

JOURNAL

Patent: WO 0104301-A 1 18-JAN-2001;
HELIX RESEARCH INSTITUTE, NORIYUKI MORIKAWA, YASUHIKO MASUHO, OSHIO
OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, YURI KAWAI
OS Homo sapiens (human)
PN WO 0104301-A/1
PD 18-JAN-2001
PR 07-JUL-2000 WO 2000JP004549
PR 08-JUL-1999 JP 99P 194179, 25-APR-2000 JP 00P 128993 PR
18-OCT-1999 US 60/159586
PI NORIYUKI MORIKAWA, YASUHIKO MASUHO, TOSHIO OTA, TAKAO ISOGAI PI
TETSUO NISHIKAWA,
PI YURI KAWAI
PC C12N15/12, C12N5/10, C12N1/15, C12N1/19, C12N1/21, C12P21/02, PC
C07K14/47
PC C07K16/18, C12Q1/02, C12Q1/68
CC
FH Key Location/Qualifiers
FT CDS (59) . (2248).
1. 2405
Location/Qualifiers
1. 2405
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 2405; DB 6; Length 2405;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCACCTCTCCCGGTTTCTGCTCTCCGCGCGTGTGAGTGTGAGGAGCTTGGGTAAT 60
DB 1 GCACCTCTCCCGGTTTCTGCTCTCCGCGCGTGTGAGTGTGAGGAGCTTGGGTAAT 60
QY 61 GGGGCTGTGCGAGGCGACGCGGCTCCCTGTAAGAGAGTCTAGCTAGAACAGAGGCG 120
DB 61 GGGGCTGTGCGAGGCGACGCGGCTCCCTGTAAGAGAGTCTAGCTAGAACAGAGGCG 120
QY 121 CTTAGGTTTTCGAGAGGAGAGTCAAGGATGTTTTCGAGCGGCTGAGACGAGTGGCC 180
DB 121 CTTAGGTTTTCGAGAGGAGAGTCAAGGATGTTTTCGAGCGGCTGAGACGAGTGGCC 180
QY 181 GATAGAGGAGAGGAGGCTCCAGTGGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
DB 181 GATAGAGGAGAGGAGGCTCCAGTGGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 241 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
DB 241 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 301 CTTTGGCGGTGAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
DB 301 CTTTGGCGGTGAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 361 TGCGCGGAGCCCGGAGAGTCCCGAGGAGGAGCTGAGGCTGAGGCTGAGGCTGAGGAGT 420
DB 361 TGCGCGGAGCCCGGAGAGTCCCGAGGAGGAGCTGAGGCTGAGGCTGAGGCTGAGGAGT 420
QY 421 GGGCCAGAGCGCGCGCGGACACCTTTCTCATTCAGCGCTCGGCGGCTTTAGTACTC 480
DB 421 GGGCCAGAGCGCGCGCGGACACCTTTCTCATTCAGCGCTCGGCGGCTTTAGTACTC 480
QY 481 AGAGCGGAGCGGAGAGTAAAGAGGCTGACAGCGGCTTCTCTAGCGGCTTTAGTACTC 540
DB 481 AGAGCGGAGCGGAGAGTAAAGAGGCTGACAGCGGCTTCTCTAGCGGCTTTAGTACTC 540
QY 541 CTGGGAGACCGGACGCGCGGACAGCGGCGAGGGAGCGCTGAGAGAGCGAGCGGCGAG 600
DB 541 CTGGGAGACCGGACGCGCGGACAGCGGCGAGGGAGCGCTGAGAGAGCGAGCGGCGAG 600
QY 601 GCGGAGAGCGGAGAGTCAAGCGCGCGGAGAGCGGCGGAGGTTTTCGCGAGAGCGAGT 660
DB 601 GCGGAGAGCGGAGAGTCAAGCGCGCGGAGAGCGGCGGAGGTTTTCGCGAGAGCGAGT 660

QY 661 CGCCAGAGGTGAGAGGAGCGCGCGCTCTGCTACCTGAGGAGTGTGGCTGCTCTCT 720
DB 661 CGCCAGAGGTGAGAGGAGCGCGCGCTCTGCTACCTGAGGAGTGTGGCTGCTCTCT 720
QY 721 CCGCGCTGCGCGAGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 780
DB 721 CCGCGCTGCGCGAGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 780
QY 781 TGCCTTTTGTGCGGACCGCGCTGCGCGCGGAGGAGTCTGAGTCTGAGTCTGAGTCTGAG 840
DB 781 TGCCTTTTGTGCGGACCGCGCTGCGCGCGGAGGAGTCTGAGTCTGAGTCTGAGTCTGAG 840
QY 841 CGGCGCGGCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 900
DB 841 CGGCGCGGCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 900
QY 901 CCGAGAGGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 960
DB 901 CCGAGAGGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 960
QY 961 TACGATTTTGTGCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1020
DB 961 TACGATTTTGTGCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1020
QY 1021 TTCCCGCCAGAGCATTAACAGACAGGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1080
DB 1021 TTCCCGCCAGAGCATTAACAGACAGGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1080
QY 1081 CCCCAAGGCTGCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1140
DB 1081 CCCCAAGGCTGCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1140
QY 1141 GTGTGTGTGTCACAGAGAGATGTATCTACCTGCGCTCCCACTTCACTCACTCACTCACTCACT 1200
DB 1141 GTGTGTGTGTCACAGAGAGATGTATCTACCTGCGCTCCCACTTCACTCACTCACTCACTCACT 1200
QY 1201 TTCCCTGTGCGGATCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1260
DB 1201 TTCCCTGTGCGGATCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1260
QY 1261 GTTCTGCGGCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1320
DB 1261 GTTCTGCGGCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1320
QY 1321 CATTTGGGAGGCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1380
DB 1321 CATTTGGGAGGCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1380
QY 1381 TAAAGTCCGCGTGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
DB 1381 TAAAGTCCGCGTGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
QY 1441 GCGCTTGGGAGGCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1500
DB 1441 GCGCTTGGGAGGCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1500
QY 1501 CATCACTTACACAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
DB 1501 CATCACTTACACAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
QY 1561 CTTCCCTTCTCTGAT 1620
DB 1561 CTTCCCTTCTCTGAT 1620
QY 1621 GGGGAGCTGTAAGGAGCATCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1680
DB 1621 GGGGAGCTGTAAGGAGCATCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1680
QY 1681 GCAATGCCATTTCTGAGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740
DB 1681 GCAATGCCATTTCTGAGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740
QY 1741 GGAATGCTTCCGAGCTGGGAGATGTTTCTTCAACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1800

Db	1741	GGAGTGTCTTCGGGCTGGGGATGTTTTCTTCAACACTGGGGAACTGTGTGGTCTGGCATGA	1800
Qy	1801	CCAAAGTTTTCTCGGCTTCCATGATCGTACTGGAGAACCTTCAAGGTGGAAAGGGGAGAA	1860
Db	1801	CCAAAGTTTTCTCGGCTTCCATGATCGTACTGGAGAACCTTCAAGGTGGAAAGGGGAGAA	1860
Qy	1861	TGTGGCCACACACGAGGTGGCAGAGGTCCTTCAGAGCCCTTGAATTTCTTACAGAGGTGAA	1920
Db	1861	TGTGGCCACACACGAGGTGGCAGAGGTCCTTCAGAGCCCTTGAATTTCTTACAGAGGTGAA	1920
Qy	1921	CGTCTATGAGTCACTGTGTCACGGGCAATGAAGGCAAGGGCTGGAAATGCAAGCCTTATGTTCT	1980
Db	1921	CGTCTATGAGTCACTGTGTCACGGGCAATGAAGGCAAGGGCTGGAAATGCAAGCCTTATGTTCT	1980
Qy	1981	GGCGTCCCCCAGGCTTTGGACCTTATGACGCTTACACCAAGTGTCTGAGAACTTGCC	2040
Db	1981	GGCGTCCCCCAGGCTTTGGACCTTATGACGCTTACACCAAGTGTCTGAGAACTTGCC	2040
Qy	2041	ACCTTATGCCCGGCGCCGATTCCTCAGAGCTTCAGAGTCTTTGGCCACACAGAACCTT	2100
Db	2041	ACCTTATGCCCGGCGCCGATTCCTCAGAGCTTCAGAGTCTTTGGCCACACAGAACCTT	2100
Qy	2101	CAAAACAGACAAAGTTGGATGGCAAAATGAGGGCTTTGACCCGACACCTGTGTGACCC	2160
Db	2101	CAAAACAGACAAAGTTGGATGGCAAAATGAGGGCTTTGACCCGACACCTGTGTGACCC	2160
Qy	2161	ACTGTACGTTTGGACACAGGCTGTAGGTCCTACCTGCCCCCTCAACAATGCGCGGTACAG	2220
Db	2161	ACTGTACGTTTGGACACAGGCTGTAGGTCCTACCTGCCCCCTCAACAATGCGCGGTACAG	2220
Qy	2221	CGCCCTCTCGGACAGAAACCTTGGAACTGTGAAATCTTCACACCTTGAGGCACTGTAGAGA	2280
Db	2221	CGCCCTCTCGGACAGAAACCTTGGAACTGTGAAATCTTCACACCTTGAGGCACTGTAGAGA	2280
Qy	2281	GGAACTCTGTGGGGTGGGGGCGGTGTGACAGGTACTGGGCGTGTAGGGATTTTCTATA	2340
Db	2281	GGAACTCTGTGGGGTGGGGGCGGTGTGACAGGTACTGGGCGTGTAGGGATTTTCTATA	2340
Qy	2341	CCAGAACTGCGGCTCACTATTTTGTAAATATGTGGCTGAGCTGATCCAGCTGTCTGTGA	2400
Db	2341	CCAGAACTGCGGCTCACTATTTTGTAAATATGTGGCTGAGCTGATCCAGCTGTCTGTGA	2400
Qy	2401	CCTAC 2405	
Db	2401	CCTAC 2405	
RESULT 2			
AK075377			
LOCUS	AK075377	2405 bp	mRNA linear PRI 03-SEP-2002
DEFINITION	Homo sapiens cDNA PSEC0067 f1s, clone NT2RP2001142, weakly similar		
	to VERR-LONG-CHAIN ACYI-COA SYNTHETASE (EC 6.2.1.1-).		
ACCESSION	AK075377		
VERSION	AK075377.1	GI:22761422	
KEYWORDS	oligo capping; f1s (full insert sequence).		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE			
AUTHORS	1 Ota,T., Nishikawa,T., Suzuki,Y., Kawai-Hio,Y., Hayashi,K.,		
	Ishii,S., Saito,K., Yamamoto,J., Wakamatsu,A., Nagai,T.,		
	Nakamura,Y., Nagahara,K., Sugano,S. and Isogai,T.		
	HRI human cDNA sequencing project		
TITLE	Unpublished		
JOURNAL	2 (bases 1 to 2405)		
REFERENCE	Isogai,T. and Yamamoto,J.		
AUTHORS	Direct Submission		
TITLE	Submitted (20-Mar-2002) Takao Isogai, Helix Research Institute,		
JOURNAL	Genomics Laboratory; 1532-3 Yana, Kizarezu Chiba 292-0812, Japan		
	(E-mail:genom@csbri1.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)		
COMMENT	HRI human cDNA sequencing project; cDNA 5'- & 3'-end one pass		

[illegible]

Db 481 AGAGGCGAGCGCGAGTTAACAAGGCTGACAGCCCTTCTCACTGCTGCTAGGCTGGA 540
 Qy 541 CTGGGGAACCGGACCGCGCGAGCAGCGGAGGAGCGCTGAGAGAGCGAGCGGAGC 600
 Db 541 CTGGGGAACCGGACCGCGCGAGCAGCGGAGGAGCGCTGAGAGAGCGAGCGGAGC 600
 Qy 601 GCGGGAACCGGAGAGTGCAGCGCGCGGAGCGGAGCGGAGAGTTTGCAGGAGCGGAGC 660
 Db 601 GCGGGAACCGGAGAGTGCAGCGCGCGGAGCGGAGCGGAGAGTTTGCAGGAGCGGAGC 660
 Qy 661 GCGGGAACCGGAGAGTGCAGCGCGCGGAGCGGAGCGGAGAGTTTGCAGGAGCGGAGC 660
 Db 661 GCGGGAACCGGAGAGTGCAGCGCGCGGAGCGGAGCGGAGAGTTTGCAGGAGCGGAGC 660
 Qy 721 CCGCAGAGGTGAGAGAGCGCGCGCTCTGTCACTGAGCAACTGTGCGCTGCTCT 720
 Db 721 CCGCAGAGGTGAGAGAGCGCGCGCTCTGTCACTGAGCAACTGTGCGCTGCTCT 720
 Qy 721 CCGCAGAGGTGAGAGAGCGCGCGCTCTGTCACTGAGCAACTGTGCGCTGCTCT 720
 Db 721 CCGCAGAGGTGAGAGAGCGCGCGCTCTGTCACTGAGCAACTGTGCGCTGCTCT 720
 Qy 781 TGCCCTTTGAGCCACCGCGCTGCGCGGCGCGCGCTGCTGAGCACTGCGAGCTGCG 840
 Db 781 TGCCCTTTGAGCCACCGCGCTGCGCGGCGCGCGCTGCTGAGCACTGCGAGCTGCG 840
 Qy 841 GCGGCGCGCGCTGCTGCTGCGCGCGAGATTTCTGAGATCTCTGAGCGCGAGCTGCG 900
 Db 841 GCGGCGCGCGCTGCTGCTGCGCGCGAGATTTCTGAGATCTCTGAGCGCGAGCTGCG 900
 Qy 901 CCGGAGAGCGATGAGGAGCTCACTGCTGAGGAGCGCGAGAGAGAGAGAGAGAGAG 960
 Db 901 CCGGAGAGCGATGAGGAGCTCACTGCTGAGGAGCGCGAGAGAGAGAGAGAGAGAG 960
 Qy 961 TAGCGATTTGCTGCTGAGATGCTGCGCTGAGAGTGGATGGAGCGAGTGCAGATCTCT 1020
 Db 961 TAGCGATTTGCTGCTGAGATGCTGCGCTGAGAGTGGATGGAGCGAGTGCAGATCTCT 1020
 Qy 1021 TTCCCGGAG 1080
 Db 1021 TTCCCGGAG 1080
 Qy 1081 CCGGAGAGCTGCTGAGATGAGTCACTGAGAGATCTGAGAGAGAGAGAGAGAGAG 1140
 Db 1081 CCGGAGAGCTGCTGAGATGAGTCACTGAGAGATCTGAGAGAGAGAGAGAGAGAG 1140
 Qy 1141 GGTGAGTGTCCACAG 1200
 Db 1141 GGTGAGTGTCCACAG 1200
 Qy 1201 TTCCCTGCTGAGGAGATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
 Db 1201 TTCCCTGCTGAGGAGATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
 Qy 1261 GTTCTGCGCTGCTGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
 Db 1261 GTTCTGCGCTGCTGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
 Qy 1321 CATTGAGGAGAGTGTGCGGATGAGTGTCAACAGAGAGAGAGAGAGAGAGAGAGAG 1380
 Db 1321 CATTGAGGAGAGTGTGCGGATGAGTGTCAACAGAGAGAGAGAGAGAGAGAGAGAG 1380
 Qy 1381 TAAAGTTCGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
 Db 1381 TAAAGTTCGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
 Qy 1441 GCGCTTGGGAG 1500
 Db 1441 GCGCTTGGGAG 1500
 Qy 1501 CATCACTACACAG 1560
 Db 1501 CATCACTACACAG 1560
 Qy 1561 CTTCCT 1620
 Db 1561 CTTCCT 1620

Db 1561 CTTCCT 1620
 Qy 1621 GGGGCACTGATGAGGAG 1680
 Db 1621 GGGGCACTGATGAGGAG 1680
 Qy 1681 GAGTCCCACTTCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
 Db 1681 GAGTCCCACTTCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
 Qy 1741 GAGTCCCACTTCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
 Db 1741 GAGTCCCACTTCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
 Qy 1801 CCAAGGTTTCT 1860
 Db 1801 CCAAGGTTTCT 1860
 Qy 1861 TGTGSCCAACAACAG 1920
 Db 1861 TGTGSCCAACAACAG 1920
 Qy 1921 CGTCTATGAGATCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980
 Db 1921 CGTCTATGAGATCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980
 Qy 1981 GGTGSCCAACAACAG 2040
 Db 1981 GGTGSCCAACAACAG 2040
 Qy 2041 ACTTATGAG 2100
 Db 2041 ACTTATGAG 2100
 Qy 2101 CAAAG 2160
 Db 2101 CAAAG 2160
 Qy 2161 ACTGATGAG 2220
 Db 2161 ACTGATGAG 2220
 Qy 2221 GCGCTCTCTGAG 2280
 Db 2221 GCGCTCTCTGAG 2280
 Qy 2281 GGAATCTGAG 2340
 Db 2281 GGAATCTGAG 2340
 Qy 2341 CCAAG 2400
 Db 2341 CCAAG 2400
 Qy 2401 CCTAC 2405
 Db 2401 CCTAC 2405

RESULT 3
 AX677203 2430 bp DNA linear PAT 27-MAR-2003
 LOCUS Sequence 396 from Patent WO02103028.
 DEFINITION AX677203
 ACCESSION AX677203.1 GI:29334629
 VERSION
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 Baranova, A.V., Yankovsky, N.K., Kozlov, A.P., Lobashev, A.V. and
 Krukovekaya, L.L.

TITLE In silico screening for phenotype-associated expressed sequences
JOURNAL Patent: WO 02103028-A 396 27-DEC-2002;
Biomedical Center (RU)

FEATURES
Source
1. 2430
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 2405; DB 6; Length 2430;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GCACTCTCCCGGGTTTCTGCTCTCCGCCGTGTGAGTGTGGGGGCTTGGGTGGAAAT 60
8 GCACTCTCCCGGGTTTCTGCTCTCCGCCGTGTGAGTGTGGGGGCTTGGGTGGAAAT 67
61 GGGGCTGTGCGACGCGACGCGCGCTCCCTGGAAGAGAAATCTCAGCTAGAAAGAGCGGC 120
68 GGGGCTGTGCGACGCGACGCGCGCTCCCTGGAAGAGAAATCTCAGCTAGAAAGAGCGGC 127
121 CCTAGCTTTTGGAGAGGAGGATCAGGAGTGTTCGAGCGGCTGAAACGACGAGTGGCC 180
128 CTTAGCTTTTGGAGAGGAGGATCAGGAGTGTTCGAGCGGCTGAAACGACGAGTGGCC 187
181 GATAGAGAGAGCGGCTTCATAGCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
188 GATAGAGAGAGCGGCTTCATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 247
241 GCT 300
248 GCT 307
301 CTTTGGCGTGTGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
308 CTTTGGCGTGTGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 367
361 TGGCGGCGACCCCGAGAGTCCCGAGAGGGGGCTGAGCTGAGCTGAGCGGCTCGGAGAACT 420
368 TGGCGGCGACCCCGAGAGTCCCGAGAGGGGGCTGAGCTGAGCTGAGCGGCTCGGAGAACT 427
421 GGGCGAGAGCGCGCGCGACACCTTCTCATTTCAAGGCTGCGGCGCTTTTAACTATC 480
428 GGGCGAGAGCGCGCGCGACACCTTCTCATTTCAAGGCTGCGGCGCTTTTAACTATC 487
481 AGAGCGGAGCGCGAGAGTAAAGAGGCTGCAAGCGGCTTCTTAAGTGTGCTAAGCTAGGA 540
488 AGAGCGGAGCGCGAGAGTAAAGAGGCTGCAAGCGGCTTCTTAAGTGTGCTAAGCTAGGA 547
541 CTGGGAGACCCGAGCGCGAGAGCGGCGAGCGGCGAGCGGCTGAGAGAGCGAGCGAGC 600
548 CTGGGAGACCCGAGCGCGAGAGCGGCGAGCGGCGAGCGGCTGAGAGAGCGAGCGAGC 607
601 GCGGAGACCCGAGAGTGAAGCGCGGAGAGCGGCGAGAGTGTTCGCGAGAGGAGTGTTC 660
608 GCGGAGACCCGAGAGTGAAGCGCGCGAGAGCGGCGAGAGTGTTCGCGAGAGGAGTGTTC 667
661 GCGGAGAGTGAAGAGCGCGCGAGAGCGGCGAGAGCGGCGAGAGTGTTCGCGAGAGGAGTGTTC 720
668 GCGGAGAGTGAAGAGCGCGCGAGAGCGGCGAGAGCGGCGAGAGTGTTCGCGAGAGGAGTGTTC 727
721 CCGCGCTGGCCGAGAGTGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
728 CCGCGCTGGCCGAGAGTGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 787
781 TGCCTTTTGGCCACCGCGCTGCG 840
788 TGCCTTTTGGCCACCGCGCTGCG 847
841 CCGCGCGCGCTGAGTGTGCG 900
848 CCGCGCGCGCTGAGTGTGCG 907

901 CCTGAGACCATGGGGCTTCACCTTGTGGCTGCAAGGCCAGAGAACCCACCTGCTGGAAT 960
908 CCGAGAGCATGGGGCTTCACCTTGTGGCTGCAAGGCCAGAGAACCCACCTGCTGGAAT 967
961 TAGGATTTTGGGCTGGAAGTCCGCTGAAGTGAAGGGCCAGTGCAGATACCTTC 1020
968 TAGGATTTTGGGCTGGAAGTCCGCTGAAGTGAAGGGCCAGTGCAGATACCTTC 1027
1021 TTCCCCCAGAGCATTAACAGACGTCCTGTATCATCTTCACTTGTGGCAACAGCGGCT 1080
1028 TTCCCCCAGAGCATTAACAGACGTCCTGTATCATCTTCACTTGTGGCAACAGCGGCT 1087
1081 CCGCAAGGCTGTGGATACGATCTGAAGATCTGTGAATGCCAGGAGCTTATCAGCT 1140
1088 CCGCAAGGCTGTGGATACGATCTGAAGATCTGTGAATGCCAGGAGCTTATCAGCT 1147
1141 GTGTGTGTCAACAGAGAGATGTATCTACTGCGCTCCACTCTACCATATGTCGG 1200
1148 GTGTGTGTCAACAGAGAGATGTATCTACTGCGCTCCACTCTACCATATGTCGG 1207
1201 TTCCCTGCTGGCATGTGTGGCTGATGGCAATTGGGGCCACAGTGTGTGAATCCAA 1260
1208 TTCCCTGCTGGCATGTGTGGCTGATGGCAATTGGGGCCACAGTGTGTGAATCCAA 1267
1261 GTTCTGAGTGTCAAGTTCTGGGAAGATTGCCAGAGACAGGTTGACGTTTCAGTA 1320
1268 GTTCTGAGTGTCAAGTTCTGGGAAGATTGCCAGAGACAGGTTGACGTTTCAGTA 1327
1321 CATTTGGAGGCTGTGTGCGATCTTGTCAACAGCCCGCGAGAGAGCGAGAACGTGGCCA 1380
1328 CATTTGGAGGCTGTGTGCGATCTTGTCAACAGCCCGCGAGAGAGCGAGAACGTGGCCA 1387
1381 TAAAGTCCGCTGGAGAGTGGAGAGCGGCTGCGCCAGATACCTGAGAGCTTTTGGG 1440
1388 TAAAGTCCGCTGGAGAGTGGAGAGCGGCTGCGCCAGATACCTGAGAGCTTTTGGG 1447
1441 GCGCTTCCGAGCCCTGCAAGTGTGTGAGACATATGAGTCAAGAGAGGCACTGGCCAC 1500
1448 GCGCTTCCGAGCCCTGCAAGTGTGTGAGACATATGAGTCAAGAGAGGCACTGGCCAC 1507
1501 CATCACTACACAGAGCAACCGGGCGCTGTGGGGGTGCTTCTGCTTTTAAAGCATAT 1560
1508 CATCACTACACAGAGCAACCGGGCGCTGTGGGGGTGCTTCTGCTTTTAAAGCATAT 1567
1561 CTTCCCTTCTCCTTGAATTCGATATGATCAACAGAGAGCCAAATCGGAGCCCGCA 1620
1568 CTTCCCTTCTCCTTGAATTCGATATGATCAACAGAGAGCCAAATCGGAGCCCGCA 1627
1621 GGGGCACTGTATGAGCCACATCTCAGGTAGACAGGGCTGTGTGGGCCCGGTAGACCA 1680
1628 GGGGCACTGTATGAGCCACATCTCAGGTAGACAGGGCTGTGTGGGCCCGGTAGACCA 1687
1681 GCACTCCCATTTCTGAGCTATGCTGAGCGGCGAGAGCTGCGCCAGGGAGATGTCTAAA 1740
1688 GCACTCCCATTTCTGAGCTATGCTGAGCGGCGAGAGCTGCGCCAGGGAGATGTCTAAA 1747
1741 GATGTCTTCCGCGCTGGGAGATTTTCTTCAACACTGGGAGCTGTGTGTGCGATGA 1800
1748 GATGTCTTCCGCGCTGGGAGATTTTCTTCAACACTGGGAGCTGTGTGTGCGATGA 1807
1801 CCAAGGTTTCTCCGCTTCCATATGATCTGAGAGACCTTCAAGTGAAGGGAGAA 1860
1808 CCAAGGTTTCTCCGCTTCCATATGATCTGAGAGACCTTCAAGTGAAGGGAGAA 1867
1861 TGTGCGCACCAACGAGTGTGAGAGTCTTGAAGCTTGAATTTCTTGAAGGTGA 1920
1868 TGTGCGCACCAACGAGTGTGAGAGTCTTGAAGCTTGAATTTCTTGAAGGTGA 1927
1921 CGTCTATGAGTCACTGTGCAAGGATGAAGGCAAGGCTGGAATGCGAGCCCTAGTTCT 1980
1928 CGTCTATGAGTCACTGTGCAAGGATGAAGGCAAGGCTGGAATGCGAGCCCTAGTTCT 1987

QY 1981 GCGTCCCGCCCAAGCTTTGAGACCTTATGACAGCTCTACACCAAGTGTCTGAGAACTTGCC 2040
 DB 1988 GCGTCCCGCCCAAGCTTTGAGACCTTATGACAGCTCTACACCAAGTGTCTGAGAACTTGCC 2047
 QY 2041 ACTTATGCGCCGCGCCCGAATCTCTCAAGGCTTCCAGAGTCTTGGCCACCAAGAGACCTT 2100
 DB 2048 ACTTATGCGCCGCGCCCGAATCTCTCAAGGCTTCCAGAGTCTTGGCCACCAAGAGACCTT 2107
 QY 2101 CAAACAGAGAAAGTTGCGATGCGCAATAGAGGCTTCCAGCCAGACCCCTGTCTGAGACC 2160
 DB 2108 CAAACAGAGAAAGTTGCGATGCGCAATAGAGGCTTCCAGCCAGACCCCTGTCTGAGACC 2167
 QY 2161 ACTGATGCTTCTGAGACAGGCTGTAGTGCCTTACCTGCCCTCAGACACTGCCCGGTACAG 2220
 DB 2168 ACTGATGCTTCTGAGACAGGCTGTAGTGCCTTACCTGCCCTCAGACACTGCCCGGTACAG 2227
 QY 2221 CGCCCTCTGAGAGAGAACTTGAATCTGAGAACTTCCACCTGAGGACCTGAGAGA 2280
 DB 2228 CGCCCTCTGAGAGAGAACTTGAATCTGAGAACTTCCACCTGAGGACCTGAGAGA 2287
 QY 2281 GGAACCTGTGTGGGGTGGGGCGGTGACAGTGTACTGTGGGCTGTGAGGATCTTTCTATA 2340
 DB 2288 GGAACCTGTGTGGGGTGGGGCGGTGACAGTGTACTGTGGGCTGTGAGGATCTTTCTATA 2347
 QY 2341 CGAAGACCTGCGCTCACTATTGTGTAATAATGTGCTGAGCTGATTCAGCTGTCTCTGA 2400
 DB 2348 CGAAGACCTGCGCTCACTATTGTGTAATAATGTGCTGAGCTGATTCAGCTGTCTCTGA 2407
 QY 2401 CCTAC 2405
 DB 2408 CCTAC 2412

RESULT 4
 LOCUS BC003041
 DEFINITION Homo sapiens solute carrier family 27 (fatty acid transporter), member 3, mRNA (cDNA clone MGC:4365 IMAGE:282875), complete cds.
 ACCESSION BC003041
 VERSION BC003041.1 GI:12804360
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 2430)
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klusner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Alech, S.F., Zeeberg, B., Bluetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F., Dietzenko, L., Marusik, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Ueda, T.B., Yoshizaki, S., Carmine, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Woley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, J., Whiting, M., Madan, A., Young, A.C., Shachenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butler, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E., Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL MEDLINE
 2288257
 12477932
 PUBLISHED 2 (bases 1 to 2430)
 REFERENCE Strausberg, R.
 AUTHORS Direct Submission
 TITLE Submitted (06-FEB-2001) National Institutes of Health, Mammalian

REMARK COMMENT
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: DCTD/DRP
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@cgsc.bc.ca
 Steven Jones, Jennifer Aano, Ian Bosdet, Yaron Butterfield,
 Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
 Leticia Hsiao, Martin Krzywinski, Rena Kutsche, Oliver Lee, Soo
 Sen Lee, Victor Ling, Carrie Matheson, Candice McLeay, Steven
 Ness, Pawan Pandon, Anna-Lisa Pradhan, Parvaneh Saeedi, Jacqueline
 Schein, Duane Smal, Michael Smith, Loraine Spence, Jeff Stott,
 Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
 George Yang, Scott Zuyderduyn, Marco Marra.

FEATURES
 source
 clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAL Plate: 9 Row: n Column: 23
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 1236578.
 Location/Qualifiers
 1..2430
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:4365 IMAGE:282875"
 /issue_type="lung, small cell carcinoma"
 /clone_lib="NIH MGC-7"
 /lab_host="DH10B-R"
 /note="vector: pOTB7"
 1..2430
 /gene="SLC27A3"
 /note="synonyms: FATP3, MGC4365"
 /db_xref="locusid:11000"
 /db_xref="MIM:604193"
 66..2258
 /codon_start=1
 /product="solute carrier family 27 member 3"
 /protein_id="P440104.1"
 /db_xref="GI:12804361"
 /translaton="MGVCGTRPAPWKEKSLERAPALGPRKSGGMPFAGMNOVPIEE
 AGSMALILPLILPLILPLILKHLMPOLRWLPDLAVAVALCKKALRALALAAA
 ADPEEGEGSGIARILAEIACQRAHATFLIGSRFSYSERASNSNAARFALGM
 DMGPRGEGSGIARILAEIACQRAHATFLIGSRFSYSERASNSNAARFALGM
 LILPAGPFLWMLFELAKAGRTAVPPLALRGLPLHLKRCGKALVALAEFLSLE
 TGLTLRPMGLHMLAAGTHTPAISIDLAEVSAEVPVPGVYSQSITDTLYTF
 LSGTGLPKARHISLKLIOCGFYOLGVAQEDVILVLAHYMSSGLIGVCMGI
 GATVILKSPFAGFQWEDCOQRTVYQYIELRVYNOPPSAEGHKKRLAVSG
 LRPDWERFVRFGPLQVLEFYELGEGVATINTGCGAVGASMTYKHFPSLR
 GVTGTEGIRDPQCNATSPREGILVAPVSOQSPILGAGPELAKQGLIKVPRP
 GVPFNRDILVCCDQGLRPHRDGTGTFPRKSNVATTVAVAFALDPELOENRVG
 VTVHEGRAMALVLRPPLALDMQVTVHSNLPVAPRFLRQESLATEYTRK
 QOKRMANEGDPSTLSPLVLDQAVGAVPLTAYVSALLAGNLRI"
 711..2147
 /note="Calc; Region: Acyl-CoA synthetases
 (AMP-forming)/AMP-acid ligases II [lipid metabolism /
 secondary metabolites biosynthesis, transport, and
 catabolism]"
 /db_xref="CDD:COG0318"

ORIGIN
 Query Match 100.0%; Score 2405; DB 9; Length 2430;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GCACTCTCCCGGGGTTTCTCTCCGCGCGCTGTGAGTGTGGGGCGCTGGGTGGGAT 60

Db	8	GCACCTCCCGGGGTTCTGCTCTCCGCCGCTGTGAGTGGTGGAGGACCTGGGAT	67
Qy	61	GGGCGTGTGACCGCACGCGCGCTCCCTGGAAAGAAAGTCTTACGTAAGACGAGCGC	120
Db	68	GGCGGTGTGACGACGACGCGCGCTCCCTGGAAAGAAAGTCTTACGTAAGACGAGCGC	127
Qy	121	CTTAGGTTTTGGAAGGAGATATAGGAGTGTTTGGAGGGGCTGGAAACGAGCGGTGC	180
Db	128	CTTAGGTTTTGGAAGGAGAGATATAGGAGTGTTTGGAGGGGCTGGAAACGAGCGGTGC	187
Qy	181	GATAGAGGAAGCGAGGCTCCATAGCTGGCCCTCTGCTGCGCCCTGCTGTGTACCC	240
Db	188	GATAGAGGAAGCGAGGCTCCATAGGCTGGCCCTCTGCTGCGCCCTGCTGTGTACCC	247
Qy	241	GCTGCTGCTGTGAAGCTTACACCTCTGGCCGCAAGTGGCTGGCTTCCGCGCACTTGGC	300
Db	248	GCTGCTGCTGTGAAGCTTACACCTCTGGCCGCAAGTGGCTGGCTTCCGCGCACTTGGC	307
Qy	301	CTTGGCGGTGGAGCTGTGCTGCACAAAGGAGCTTTCGAGCTGCGCCCTGGCGCGCGC	360
Db	308	CTTGGCGGTGGAGCTGTGCTGCACAAAGGAGCTTTCGAGCTGCGCCCTGGCGCGCGC	367
Qy	361	TGCCGCGACCCCGAAGGTCCCGAGGGGGGCTGACGCTTGCGCTGCGCGAAT	420
Db	368	TGCCGCGCGACCCCGAAGGTCCCGAGGGGGGCTGACGCTTGCGCTGCGCGAAT	427
Qy	421	GGCCCGAGAGGGCCCGCGCACCTTCTTCATTCAGGGCTCGGCGGGCTTTAGCTACCTC	480
Db	428	GGCCCGAGAGGGCCCGCGCACCTTCTTCATTCAGGGCTCGGCGGGCTTTAGCTACCTC	487
Qy	481	AGAGCGGAGCGCGAGAGTAAACAGGGCTGCACGCGCTTCTTACGTGCGTAAAGCTGGGA	540
Db	488	AGAGCGGAGCGCGAGAGTAAACAGGGCTGCACGCGCTTCTTACGTGCGTAAAGCTGGGA	547
Qy	541	CTGGGGACCCGACCGCGCGCGACAGCGGCTGAAGGGACGCTGGAGAAGCCGAGCGCGAC	600
Db	548	CTGGGGACCCGACCGCGCGCGACAGCGGCTGAAGGGACGCTGGAGAAGCCGAGCGCGAC	607
Qy	601	GCCGGGACCCGGAAGTTCAGCGCGCGAGGCGCGGAGGTTTCCGGAAGGAGCGAGTGC	660
Db	608	GCCGGGACCCGGAAGTTCAGCGCGCGAGGCGCGGAGGTTTCCGGAAGGAGCGAGTGC	667
Qy	661	GCCGAGAGTGAAGAGCGCGCGCTCTGTGTACCTTGAGCAACCTGTGCTGCTCT	720
Db	668	GCCGAGAGTGAAGAGCGCGCGCTCTGTGTACCTTGAGCAACCTGTGCTGCTCT	727
Qy	721	CCCCGCTGAGCCAGGTTTCGTGGCTCTGGTTGGGGTGGGCAAGCGCGGCTGGCGAC	780
Db	728	CCCCGCTGAGCCAGGTTTCGTGGCTCTGGTTGGGGTGGGCAAGCGCGGCTGGCGAC	787
Qy	781	TGCTCTTTGCGCACCGCGCTGCGCGGGGCGCCCTGCTGCACTGCTTCGCGAGCTGCG	840
Db	788	TGCTCTTTGCGCACCGCGCTGCGCGGGGCGCCCTGCTGCACTGCTTCGCGAGCTGCG	847
Qy	841	GCGCGCGCGCTGTGTGCTGGCGCCAGATTTCTGGAATCCTTGAGCCCGACCTGCGCGC	900
Db	848	GCGCGCGCGCTGTGTGCTGGCGCCAGATTTCTGGAATCCTTGAGCCCGACCTGCGCGC	907
Qy	901	CTTGAGAGCCATGAGGAGCTTCACTTGTGGGCTTGACAGCCACAGAAACCACTTGTGAAAT	960
Db	908	CTTGAGAGCCATGAGGAGCTTCACTTGTGGGCTTGACAGCCACAGAAACCACTTGTGAAAT	967
Qy	961	TAGCGATTTTGCTGCTGAAGTGTCCGCTGAAGTGAATGGCGCAGTGCAGAGTACCTTC	1020
Db	968	TAGCGATTTTGCTGCTGAAGTGTCCGCTGAAGTGAATGGCGCAGTGCAGAGTACCTTC	1027
Qy	1021	TTCCCCCGAGACATTAACAGACACGTGCTGTACATCTTCACTTGTGCAACAAGGAGCT	1080
Db	1028	TTCCCCCGAGACATTAACAGACACGTGCTGTACATCTTCACTTGTGCAACAAGGAGCT	1087
Qy	1081	CCCGAAGGCTGCTCGATCAATCATGGAAGTCTTGCAATGCCAGGCTTCTATCAGCT	1140

Db	1088	CCCCAAG3CTGCTCGAGTACAGTACATCTGAAGATCTTGCAATGCGAGGCGCTTTCATACGCT	1144
Qy	1141	GTGTGTGTGTCCACAGAAAGATGTGATCTACCTCGCCCTCCACCTACACCAATGTCCGG	1200
Db	1148	GTGTGTGTGTCCACAGAAAGATGTGATCTACCTCGCCCTCCACCTACACCAATGTCCGG	1207
Qy	1201	TTCCCTCTGGGACATCGTGGGGCTGCATGGGACATTTGGGGCCACAGTGTGTGTAATCA	1260
Db	1208	TTCCCTCTGGGACATCGTGGGGCTGCATGGGACATTTGGGGCCACAGTGTGTGTAATCA	1267
Qy	1261	GTTCTCGGCTGTGATTTCTGGGAAGATTGCCAGACACAAAGGTGA CGGTGTTCCAGTA	1320
Db	1268	GTTCTCGGCTGTGATTTCTGGGAAGATTGCCAGACACAAAGGTGA CGGTGTTCCAGTA	1327
Qy	1321	CATTGGGGAGGTGTGCCGATACCTTTGTCAACACGAGCCCGGAGCAAGGCAACGATGGCCA	1380
Db	1328	CATTGGGGAGGTGTGCCGATACCTTTGTCAACACGAGCCCGGAGCAAGGCAACGATGGCCA	1387
Qy	1381	TAAAGTCCGCTGAGCATGTG3CAGCGGCGCTGTGGCGCCAGATACCTGGAGCGTTTGTGCG	1440
Db	1388	TAAAGTCCGCTGAGCATGTG3CAGCGGCGCTGTGGCGCCAGATACCTGGAGCGTTTGTGCG	1447
Qy	1441	GCGCTTTGGGGCCCGCTGAGGTCGTGAGACATTA TGA CTGACACAGAGG3CAACGTG3CAC	1500
Db	1448	GCGCTTTGGGGCCCGCTGAGGTCGTGAGACATTA TGA CTGACAGAGG3CAACGTG3CAC	1507
Qy	1501	CATCAACTACACAGAGACAGCGG3CGCTGTGGGCGGTGCTTCTGCTTACAAAGCATAT	1560
Db	1508	CATCAACTACACAGAGACAGCGG3CGCTGTGGGCGGTGCTTCTGCTTACAAAGCATAT	1567
Qy	1561	CTTCCCTCTTCTCTGTGATTCGCTATGATATGTCACACACAGAGAGCGCATTTGGGACCCCA	1620
Db	1568	CTTCCCTCTTCTCTGTGATTCGCTATGATATGTCACACAGAGAGCGCATTTGGGACCCCA	1627
Qy	1621	GGGGACCTGTATG3CACACATCTCCAGGCTGAGCCAGG3CTGCTGGTGGCCCGCGTAAAGCA	1680
Db	1628	GGGGACCTGTATG3CACACATCTCCAGGCTGAGCCAGG3CTGCTGGTGGCCCGCGTAAAGCA	1687
Qy	1681	GCAGTCCCCATTCCTGG3CTATGCTGTGGCGGGCCAGAGCTGGCCCAAGGGGAAAGTTGCTTAA	1740
Db	1688	GCAGTCCCCATTCCTGG3CTATGCTGTGGCGGGCCAGAGCTGGCCCAAGGGGAAAGTTGCTTAA	1747
Qy	1741	GGATGTCTTCGGGCGCTGGGAGATGTTTTCTTCAACATCGGG3ACCTGCTGATCTCCAGTGA	1800
Db	1748	GGATGTCTTCGGGCGCTGGGAGATGTTTTCTTCAACATCTGGG3ACCTGCTGATCTCCAGTGA	1807
Qy	1801	CCAAAGTTTCTCCGCTTCCATGATGCTGATCTGAGACACCTTCAGGTGAAAGG3G3AGAA	1860
Db	1808	CCAAAGTTTCTCCGCTTCCATGATGCTGATCTGAGACACCTTCAGGTGAAAGG3G3AGAA	1867
Qy	1861	TGTGGCCACACCGAGTGTGAGAGAGTCTTGAGAGCCCTAGATTTCTTCAGAGAGTGAA	1920
Db	1868	TGTGGCCACACCGAGTGTGAGAGAGTCTTGAGAGCCCTAGATTTCTTCAGAGAGTGAA	1927
Qy	1921	CGTCTATGAGATCTATGTGCCAGG3CATGAAAGCGAG3CTGGAATGCGACGCTAGTTCT	1980
Db	1928	CGTCTATGAGATCTATGTGCCAGG3CATGAAAGCGAG3CTGGAATGCGACGCTAGTTCT	1987
Qy	1981	GCGTCCCCCCCCAGCTTTTGAACCTTATGCACTTACACCCACGCTGTGTGAACCTTGGC	2040
Db	1988	GCGTCCCCCCCCAGCTTTTGAACCTTATGCACTTACACCCACGCTGTGTGAACCTTGGC	2047
Qy	2041	ACCTTATCCCGGCGCCGATTCCTCAGG3CTCCAGAGATCTTTTGGCCACCAAGAGACCTT	2100
Db	2048	ACCTTATCCCGGCGCCGATTCCTCAGG3CTCCAGAGATCTTTTGGCCACCAAGAGACCTT	2107
Qy	2101	CAACACAGAAAGTTCCGATGTGCAATGAGG3CTTGACCCCAAGACACCTGTGTGACCC	2160
Db	2108	CAACACAGAAAGTTCCGATGTGCAATGAGG3CTTGACCCCAAGACACCTGTGTGACCC	2167
Qy	2161	ACTGTAGTCTTGAGACCAAGGCTGTATGAGGCTTACCTG3CTTCAACATGTGCCG3TACAG	2220
Db	2168	ACTGTAGTCTTGAGACCAAGGCTGTATGAGGCTTACCTG3CTTCAACATGTGCCG3TACAG	2227

Qy	2221	GGCCCTCCGCGGAGGAAACCTTCGAACTGAGAACTTCACAACCCGAGGACACTGAGGA	2280
Db	2228	CGCCCTCTCGGCGAGGAAACCTTCGAACTGAGAACTTCACAACCCGAGGACACTGAGGA	2287
Qy	2281	GGAACTCTGTGGGAGTGGGCGCGGTGACAGGTGACTGCGGCTGTCAAGGATCTTTTCTATA	2340
Db	2288	GGAACTCTGTGGGAGTGGGCGCGGTGACAGGTGACTGCGGCTGTCAAGGATCTTTTCTATA	2347
Qy	2341	CCAGAACTGCGGCTCACTATTTTGTATATAAAGTGGCTGAGCTGATTCAGAGTGTCTCGA	2400
Db	2348	CCAGAACTGCGGCTCACTATTTTGTATATAAAGTGGCTGAGCTGATTCAGAGTGTCTCTCGA	2407
Qy	2401	CTTAC 2405	
Db	2408	CTTAC 2412	
RESULT 5			
LOCUS	BC009916	2430 bp	mRNA
DEFINITION	Homo sapiens solute carrier family 27 (fatty acid transporter), member 3, mRNA (CDNA clone MGC:2739 IMAGE:2822875), complete cds.		
ACCESSION	BC009916		
VERSION	BC009916.2	GI:33869756	
KEYWORDS	MGC.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 2430)		
	Straussberg, R. L., Feingold, E. A., Grouse, L. H., Derge, J. G., Klausner, R. D., Collins, F. S., Wagner, L., Shenmen, C. M., Schuler, G. D., Altschul, S. F., Zeeberg, B., Buetow, K. H., Schaefer, C. F., Bhat, N. K., Hopkins, R. F., Jordan, H., Moore, T., Max, S. I., Wang, J., Heide, F., Dietzen, R. J., Marusina, K., Farmer, A. A., Rubin, G. M., Hong, L., Slepenon, M., Soares, M. B., Bonaldo, M. F., Casavant, T. L., Scheer, T. E., Brownstein, M. J., Udell, T. B., Toshiyuki, S., Scarinci, P., Prange, C., Raha, S. S., Loggellano, N. A., Peters, G. J., Abramson, R. D., Mullahy, S. J., Bosak, S. A., McEwan, P. J., McGernan, K. U., Malek, V. A., Gunaratne, P. H., Richards, S., Wotley, K. C., Hale, S., Garcia, A. M., Gay, L. U., Hulyk, S. W., Villalón, D. K., Muzny, D. M., Sodergren, E. J., Lu, X., Gibbs, R. A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A. C., Shigichenko, Y., Bouffard, G. G., Blakesley, R. W., Touchman, J. W., Green, E. D., Dickson, M. C., Rodriguez, A. C., Grimwood, J., Schmutz, J., Myers, R. M., Butlerfield, Y. S., Krzyzinski, M. I., Skalska, U., Smalins, D. E., Scherch, A., Schein, J. B., Jones, S. U., and Marz, M. A.		
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)		
MEDLINE	22388257		
PUBMED	12477932		
REFERENCE	2 (bases 1 to 2430)		
AUTHORS	Straussberg, R.		
TITLE	Direct Submission		
JOURNAL	Submitted (02-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov		
COMMENT	On Aug 19, 2003 this sequence version replaced gi:14602825.		
	Contact: MGC help desk		
	Email: gcgaps-remail.nih.gov		
	Tissue Procurement: DCTD/DTF		
	cDNA Library Preparation: Rubin Laboratory		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLML)		
	DNA Sequencing by: Institute for Systems Biology		
	http://www.systemsbiology.org		
	contact: amadan@systemsbiology.org		
	Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whiting		

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINt at: <http://image.llnl.gov>

Series: IRAL Plate: 2 Row: P Column: 9

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA bc: 13236578.

Location/Qualifiers

```

1. .2430
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="MGC:2739 IMAGE:2822875"
    /issue_type="Lung, small cell carcinoma"
    /clone_id="NH_MGC_7"
    /lab_host="DH10B-R"
    /note="Vector: pOTB7"

1. .2430
    /gene="SLC27A3"
    /note="Synonyms: FATP3, MGC4365"
    /db_xref="LOCusID:11000"
    /db_xref="MIM:604193"
    .66. .2258
    /codon_start=1
    /product="solute carrier family 27 member 3"
    /protein_id="AAH09916.1"
    /db_xref="GI:14602826"
    /db_xref="LOCusID:11000"
    /translation="MGVCQTRAPWKKSKQLDEALALCFKSGSGMPSGNCQVPIIE
AGSVALLLPLPLPLPLPLKULMFQULWLPADLDAFARALCCRALPARLALAA
ADPEGDEGSGSLAWRLLEIAQORAHFTHHSRREYSBSAESNNAAALPRAICM
DMGDDGDSGSGSGBERBAFGAGAAAGAEFAGDDAARGGAAALPSPGATVP
LLLPAGEPEPLMLWFGALAGLRTAFVPTLRRPPLHCLSCCARALVLPBELSE
PDLPALARAGLHLMAAPGTHPAIGISDLIAEVAEDGVPVAGISPSQSIDPCLVF
TSGTGTGPKAARISHKTILOCGFPYOLCGHOEDVITYALPLVHMSGSLIGYCMGI
GATVYKSKRSAGOFEMEDCOOHRTVYVOYIGELCRVYVNPBPKARGHVRLAVSG
LRPWTMRPVRRRRPPDLYETXGTEBNVATTNYTGORGAVGASMLYKPIPPSLIR
YDVATGPIRDPOGHCAISPGSEGLVAESQSPFLGAGPELMOGLKQVPR
GDVFENTGDLVCDQDGFERHDTGDTPEFWSKBNVATTEVAEVELADLOKVNVG
VTVFGHGRAGAAALVLRPHALDLMLOLYHVBENLPVAPRPFRLQESLATTEFK
QOKRYANEGFDPSTLDDPLVYDQAVGALPLTTRYASALLAGNLRI"
711. .1931
    /note="AMP-binding; Region: AMP-binding enzyme"
    /db_xref="CD:pfam00501"

```

misc_feature

```

Query Match      100.0%; Score 2405; DB 9; Length 2430;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

ORIGIN	Query Match	Best Local Similarity	Matches 2405; Conservative	Score 2405; DB 9; Length 2430;	Pred. No. 0;	Mismatches 0;	Indels 0;	Gaps 0;
Db	1	GCACTCTCCCGGGTTCTGCTCTCCGCCCGCTGTGAGTGGTGGGGCGTGGGGAAT	60					
Db	8	GCACTCTCTCCCGGGTTCTGCTCTCCGCCCGCTGTGAGTGGTGGGGCGTGGGGAAT	67					
Dy	61	GGGCGTGTGCACGGCACGCGCGCTCCCTGGAAGGAAGTCTCAGCTAGACGAGCGCG	120					
Db	68	GGGCGTGTGCACGGCACGCGCGCTCCCTGGAAGGAAGTCTCAGCTAGACGAGCGCG	127					
Dy	121	CTTAGGTTTTCGGAAGGAGGAGATCAAGGATGTTTTCGAGCGGCGCTGGAAACGAGCGGTGCC	180					
Db	128	CTTAGGTTTTCGGAAGGAGGAGATCAAGGATGTTTTCGAGCGGCGCTGGAAACGAGCGGTGCC	187					
Dy	181	GATAGAGGAAGCGGCGCTCCATGCGCTGCGCTCTGCTGCTGCGCTCGCTGCTGTTGCTTAC	240					
Db	188	GATAGAGGAAGCGGCGCTCCATGCGCTGCGCTCTGCTGCTGCGCTCGCTGCTGTTGCTTAC	247					
Dy	241	GCTGCTGCTCTGTAAGCTACACTCTGAGCGGACGAGTGGAGTGGCTTCCGCGGAGACTTGGC	300					
Db	248	GCTGCTGCTCTGTAAGCTACACTCTGAGCGGACGAGTGGAGTGGCTTCCGCGGAGACTTGGC	307					
Dy	301	CTTTGCGGATGCGAGCTCTGTGCTGCAAAAGGCTCTTTGAGCTGCGCGCTTGGCGCGCGC	360					
Db	308	CTTTGCGGATGCGAGCTCTGTGCTGCAAAAGGCTCTTTGAGCTGCGCGCTTGGCGCGCGC	367					

OY	361	TGCGGCGGACCCGGGAAGTCCCGAGAGGGGGGTGGAGCTGAGCTGAGCTCGGCGGAACT	420
Db	368	TGCGCGCGACCCGGGAAGTCCCGAGAGGGGGGTGGAGCTGAGCTGAGCTCGGCGGAACT	427
OY	421	GGCCCAACAGCGCGCCCGGCAACACTTTCATTTCAAGGCTCGCGGCGCTTTAGACTC	480
Db	428	GGCCCAACAGCGCGCCCGGCAACACTTTCATTTCAAGGCTCGCGGCGCTTTAGACTC	487
OY	481	AGAGGCGGAGCGCGAGAGTAAACAGGCTGCACGGGCTTCTACGTGCGCTAGCGTGGGA	540
Db	488	AGAGGCGGAGCGCGAGAGTAAACAGGCTGCACGGGCTTCTACGTGCGCTAGCGTGGGA	547
OY	541	CTGGGGAACCCCAACGCGGCGCAACAGCGCGAGCGCTGAGAGAACGCAAGCGGCGAGC	600
Db	548	CTGGGGAACCCCAACGCGGCGCAACAGCGCGAGCGCGCTGAGAGAACGCAAGCGGCGAGC	607
OY	601	GCCGCGAGCCGGAATGACGCGCGCGAGCGCGCTGGAGATTTCGGAGGGGAGCGGTGC	660
Db	608	GCCGCGAGCCGGAATGACGCGCGCGAGCGCGCTGGAGATTTCGGAGGGGAGCGGTGC	667
OY	661	CGCCAGAGAGTGAAGAGCGCGCGGCGCTGTCTACCTGGAGCAACTGTGGCGGTCTCT	720
Db	668	CGCCAGAGAGTGAAGAGCGCGCGGCGCTGTCTACCTGGAGCAACTGTGGCGGTCTCT	727
OY	721	CCCGGCTGGCCAGAGTTTCTGTGAGCTCTGGATTGGGCTGGCCAGAGCCGGCTTGCGAC	780
Db	728	CCCGGCTGGCCAGAGTTTCTGTGAGCTCTGGATTGGGCTGGCCAGAGCCGGCTTGCGAC	787
OY	781	TGCGCTTTGTGCCACCGCGCTGGCGCGCGGGGGCGCCCTGTGCACGTGCGGACGTGCGG	840
Db	788	TGCGCTTTGTGCCACCGCGCTGGCGCGCGGGGGCGCCCTGTGTGCACGTGCGGACGTGCGG	847
OY	841	CGCGCGCGCGCTGGGTGCGCGCGCGAGATTCTGGAAGTCCCTGGAGCGCGACCTGCCGC	900
Db	848	CGCGCGCGCGCTGGGTGCGCGCGCGAGATTCTGGAAGTCCCTGGAGCGCGACCTGCCGC	907
OY	901	CCTGAGAGCCATGGGGCTCCACCTGTGGGGCTGCAGGCGCCAGAAACCAACTGCTGGAAAT	960
Db	908	CCTGAGAGCCATGGGGCTCCACCTGTGGGGCTGCAGGCGCCAGAAACCAACTGCTGGAAAT	967
OY	961	TAGGCAATTGCTGGGTGAAGTGTCCGCTGGAAGTGGATGGGCGAGGCCAGGATACCTCTC	1020
Db	968	TAGGCAATTGCTGGGTGAAGTGTCCGCTGGAAGTGGATGGGCGAGGCCAGGATACCTCTC	1027
OY	1021	TTCCGCCAGAGCATPACAGACACGTGCTGTACATCTTCACTCTGTGGACCAACGGGCTT	1080
Db	1028	TTCCGCCAGAGCATPACAGACACGTGCTGTACATCTTCACTCTGTGGACCAACGGGCTT	1087
OY	1081	CCCGAAGGCTCTCGGATCAGTCATCTGAATCTGTGAATCTGTGAATGCCAGGGCTTCTATACCT	1140
Db	1088	CCCGAAGGCTCTCGGATCAGTCATCTGAATCTGTGAATCTGTGAATGCCAGGGCTTCTATACCT	1147
OY	1141	GTGTGTGTCCACACAGAAAGATGTATCTACTCGGCTCTCCACTCTACCAATGTCCGG	1200
Db	1148	GTGTGTGTCCACACAGAAAGATGTATCTACTCGGCTCTCCACTCTACCAATGTCCGG	1207
OY	1201	TTCCCTGCTGGGCACTGTGGGCTGCATGGGCAATTGGGGCCACAGTGTGTCTGAATCCAA	1260
Db	1208	TTCCCTGCTGGGCACTGTGGGCTGCATGGGCAATTGGGGCCACAGTGTGTCTGAATCCAA	1267
OY	1261	GTTCCTCGGCTGTAGTTCTGGGAAGATTGCGACAGACACAGGGTGAAGGTGTCCAGTA	1320
Db	1268	GTTCCTCGGCTGTAGTTCTGGGAAGATTGCGACAGACACAGGGTGAAGGTGTCCAGTA	1327
OY	1321	CATTGGGAGCTGTGCGCATCTTGTCAACAGCCCTCCAGACAAAGGCAAGAGTGTGCA	1380
Db	1328	CATTGGGAGCTGTGCGCATCTTGTCAACAGCCCTCCAGACAAAGGCAAGAGTGTGCA	1387
OY	1381	TAAAGTCGCGCTGCGAGTGGGAGCGGGGCTGGCCCAATACCTGGAGCGCTTTGTGCG	1440
Db	1388	TAAAGTCGCGCTGCGAGTGGGAGCGGGGCTGGCCCAATACCTGGAGCGCTTTGTGCG	1447
OY	1441	GCGCTTGGGCGCTTGCAGAGTCTGGAACATATGAATGACACAGGGCAAGTGGCCAC	1500

Db	1448	GGCGTTTGGGGCCCCCTGAGAGTGTCTGGAGACATATGACTGACAGAGGGGCAACGTGGCCAC	1507
Qy	1501	CATCACTACACAGAGCAGCGGGCGCTGTGGGGCGTGCTTCTGTGCTTTACAAACATAT	1560
Db	1508	CATCACTACACAGAGCAGCGGGCGCTGTGGGGCGTGCTTCTGTGCTTTACAAACATAT	1567
Qy	1561	CTTCCCCCTTCTCTGTGATGTGGCTATGATATGTCACACAGAGAGAGCAATTTGGAGCCCCCA	1620
Db	1568	CTTCCCCCTTCTCTGTGATGTGGCTATGATATGTCACACAGAGAGAGCAATTTGGAGCCCCCA	1627
Qy	1621	GGGGCACTGATAGGCGCATCTCCAGGTGAGCCAGGGCTGTGGTGGCCCCGGTAAAGCCA	1680
Db	1628	GGGGCACTGATAGGCGCATCTCCAGGTGAGCCAGGGCTGTGGTGGCCCCGGTAAAGCCA	1687
Qy	1681	GCAATCCCCCATTTCTGGGGCTATGTGTGGCGGGGCCAGACTGTGGCCCAAGGGGAAAGTGGTAA	1744
Db	1688	GCAATCCCCCATTTCTGGGGCTATGTGTGGCGGGGCCAGACTGTGGCCCAAGGGGAAAGTGGTAA	1747
Qy	1741	GGATGTCTTCGGGGCTGTGGGATGTCTTTCACACACTGGGGACCTGTGTGTCTGTCGATGA	1800
Db	1748	GGATGTCTTCGGGGCTGTGGGATGTCTTTCACACACTGGGGACCTGTGTGTGTCTGTCGATGA	1807
Qy	1801	CCAAAGTTTTCCTCCGCTCCATGATGTGTACTGGAGACCTTTCAAGGTGGAAGGGGGAGAA	1866
Db	1808	CCAAAGTTTTCCTCCGCTCCATGATGTGTACTGGAGACCTTTCAAGGTGGAAGGGGGAGAA	1867
Qy	1861	TGTGGGCCAACAGAGGTGGCAGAGAGTCTTGAAGGCCCTAGATATTTCTTCAGAGAGTGA	1920
Db	1868	TGTGGGCCAACAGAGGTGGCAGAGAGTCTTGAAGGCCCTAGATATTTCTTCAGAGAGTGA	1922
Qy	1921	CGTCTATGAGATCACTGTGCCAGGGGCATGGAAGCAGGGCTGGAATGTGAGCTTATGTTCT	1988
Db	1928	CGTCTATGAGATCACTGTGCCAGGGGCATGGAAGCAGGGCTGGAATGTGAGCTTATGTTCT	1987
Qy	1981	GGTGTCCCCCCCAGGCTTTGAGACTTATGACACTTACACCCAGGTGTGAGAACTTGTGC	2040
Db	1988	GGTGTCCCCCCCAGGCTTTGAGACTTATGACACTTACACCCAGGTGTGAGAACTTGTGC	2047
Qy	2041	ACCTTATGCCCCGGGCCCGATTCCTCAGGCTTCAGAGAGTCTTTTGGCCACACAGAGACCTT	2100
Db	2048	ACCTTATGCCCCGGGCCCGATTCCTCAGGCTTCAGAGAGTCTTTTGGCCACACAGAGACCTT	2107
Qy	2101	CAAAACAGCAAAAGTTGGATGAGTGGCAAATGAGAGGCTTGACCCCCAGACCCCTGTCTGACCC	2166
Db	2108	CAAAACAGCAAAAGTTGGATGAGTGGCAAATGAGAGGCTTGACCCCCAGACCCCTGTCTGACCC	2167
Qy	2161	ACTGTAGCTTTGAGACCAAGCGCTGATAGTGCCTACCTGTGCCCTTCACATGTGCCCGGTACG	2220
Db	2168	ACTGTAGCTTTGAGACCAAGCGCTGATAGTGCCTACCTGTGCCCTTCACATGTGCCCGGTACG	2227
Qy	2221	CGCCCTCTGCGCAGAGAACTTGCATGTAGAACTTCCACACTCTGAGGCACTTGAGAG	2288
Db	2228	CGCCCTCTGCGCAGAGAACTTGCATGTAGAACTTCCACACTCTGAGGCACTTGAGAG	2287
Qy	2281	GGAACTCTGTGGGGTGGGGGCGGTTGACAGGTGTACTGTGGCTGTCAAGGATCTTTCTATA	2344
Db	2288	GGAACTCTGTGGGGTGGGGGCGGTTGACAGGTGTACTGTGGCTGTCAAGGATCTTTCTATA	2347
Qy	2341	CCAGAACTGCGGTCACTATTTTGTATATAATGTGGCTGGAAGTATCCAGCTGTCTCTGA	2400
Db	2348	CCAGAACTGCGGTCACTATTTTGTATATAATGTGGCTGGAAGTATCCAGCTGTCTCTGA	2407
Qy	2401	CTTAC 2405	
Db	2408	CTTAC 2412	

RESULT 6			
AX101424			
LOCUS	AX101424	2402 bp	DNA
DEFINITION	Sequence 116 from Patent WO021795.		
ACCESSION	AX101424		
			linear
			(PAT 10-APR-2001)

PAT 10-APR-2001

QY	1338	TGCAGGGGCATGGAAGGCAGAGGCTGGAAATGGAGACCTTGTTCTGGTCCGCCCCACGCTT	1997
Db	1994	TGCCAGGGCATTGAAGCAGGGGCTGGAAATGGCAGCCCTTGTTCTGGTCCGCCCCACGCTT	2053
QY	1998	TGACCTTATGCAGCTCTTACACCACTGTGTGAGAACCTTGCCACTTATATGCCCGGGCCCC	2057
Db	2054	TGGACCTTATGCAGCTCTTACACCACTGTGTGAGAACCTTGCCACTTATATGCCCGGGCCCC	2113
QY	2058	GATTCCTCAGGCTCCAGGAGTCTTTGGGCACACAGAGACCTTCAACAGCAGAAAGTTC	2117
Db	2114	GATTCCTCAGGCTCCAGGAGTCTTTGGGCACACAGAGACCTTCAACAGCAGAAAGTTC	2173
QY	2118	GGATGGCAATGAGGGCTTGGACCCCGACACCTGTCTGACCCACTGTAGCTTGTGAC	2177
Db	2174	GGATGGCAATGAGGGCTTGGACCCCGACACCTGTCTGACCCACTGTAGCTTGTGAC	2233
QY	2178	AGGCTGTAGGTGGCTCACTGCCCCCTCAAACTGCCCGGTACAGGCGCCCTCTGGAGAGAA	2237
Db	2234	AGGCTGTAGGTGGCTCACTGCCCCCTCAAACTGCCCGGTACAGGCGCCCTCTGGAGAGAA	2293
QY	2238	ACCTTCGATGTGGAACCTTCCACACTGAGAGCACTGAGAGAGGAATCTGTGGGGTGG	2297
Db	2294	ACCTTCGATGTGGAACCTTCCACACTGAGAGCACTGAGAGAGGAATCTGTGGGGTGG	2353
QY	2298	GGGCGCTTGCAGGTGTACTGGGCTGTCAAGGATCTTTTCTATACCAAGCTGCGGTACT	2357
Db	2354	GGGCGCTTGCAGGTGTACTGGGCTGTCAAGGATCTTTTCTATACCAAGCTGCGGTACT	2413
QY	2358	ATTTTGTATAAATGTGGCTGAGTGATTCAGAGCTGTCTGAACTTA 2404	
Db	2414	ATTTTGTATAAATGTGGCTGAGTGATTCAGAGCTGTCTGTAACTTA 2460	

RESULT 9	343 bp	mRNA	linear	PRI 09-SEP-2003
AK125102				
LOCUS				
DEFINITION	AK125102	Homo sapiens CDNA FLJ43112 f1s, clone CR2082028124, weakly similar		
ACCESSION	AK125102	to Homo sapiens VLCS-H1 protein (VLCS-H1).		
VERSION	AK125102.1	GI:34531083		
KEYWORDS		oligo capping; f1s (full insert sequence).		
SOURCE	Homo sapiens	(human)		
ORGANISM	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1			
AUTHORS	Kanehori, K., Ishibashi, T., Chiba, Y., Fujimori, K., Hiraoka, S., Tanai, H., Watanabe, S., Ishida, S., Ono, Y., Hotta, T., Watanabe, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishi, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Matsumura, M., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Matsuo, Y., Nagai, K., and Isega, T.			
TITLE	NEDO human cDNA sequencing project			
JOURNAL	unpublished			
REFERENCE	2 (bases 1 to 343)			
AUTHORS	Isega, T. and Yamamoto, J.			
TITLE	Direct Submission			
JOURNAL	Submitted (15-JUN-2003) Takao, Isega, FLJ Project (HRI Team); 2-6-7			

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clonename="CTONG2028124"
/tissue_type="conque, tumor tissue"
/clonelib="CTONG3"
/notes="cloning vector: pME18SFL1"
761..2663

```

ORIGIN

Query Match	99.3%;	Score 2387;	DB 9;	Length 3343;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2398;	Conservative 0;	Mismatches 0;	Indels 1;	Gaps 1;

QY	1	GCACCTCCTCCCGGGATTTCGTCTCTCCGCCCGGTGTGAGTGTGGGGGGCTTGGGTGGAAAT	60
Db	946	GCACCTCCTCCCGGGATTTCGTCTCTCCGCCCGGTGTGAGTGTGGGGGGCTTGGGTGGAAAT	1005
QY	61	GGGCGTGTGGCAGCGGACGCGCGCGTCCCTCGTAAGGAATACTCAGCTAGAACGAGCGGC	120
Db	1006	GGGCGTGTGGCAGCGGACGCGCGCGTCCCTCGTAAGGAATACTCAGCTAGAACGAGCGGC	1065
QY	121	CTTAGGTTTTCCGAGGAGGAGATCAGGAGATGTTTCCAGCGCGCTGGAACCAAGACGGTGC	180
Db	1066	CTTAGGTTTTCCGAGAGGAGGAGATCAGGAGATGTTTCCAGCGCGCTGGAACCAAGAGGGTGC	1125
QY	181	GATAGAGAGAGCGGGGCTCCATGGCTGCGCTCCCTGCTGCTGCTGCCCTGTGTGCGTGTGCTACC	240
Db	1126	GATAGAGAGAGCGGGGCTCCATGGCTGCGCTCCCTGCTGCTGCTGCCCTGTGTGCTGCTACC	1185
QY	241	GCTGCTGCTGCTGTAAGCTTACCTCTTGAGCCGCGAATTGCGCTTGCCGCGGACTTTGGC	300
Db	1186	GCTGCTGCTGCTGTAAGCTTACCTCTTGAGCCGCGAATTGCGCTTGCGGCTTCCGGCGGACTTTGGC	1245
QY	301	CTTTGGCGGTGCGAGCTCTGTGTGCTGCAAAAGGGGCTCTTGAGGCTGCGGCGCTCGGCGGGGC	360
Db	1246	CTTTGGCGGTGCGAGCTCTGTGTGCTGCAAAAGGGGCTCTTGAGGCTGCGGCTTGGCGGGGC	1305
QY	361	TGCCGCGACCCGGAAGGTCCCGAGGGGGGCTGCAAGCCTTGCGCTTGCGGAACT	420
Db	1306	TGCCGCGACCCGGAAGGTCCCGAGGGGGGCTGCAAGCCTTGCGCTTGCGGAACT	1365
QY	421	GGCCCAAGCGCGCGCCGCGGACACTTTTCTCATTTCAAGCGCTCGGGGGCTTTAGCTACTC	480
Db	1366	GGCCCAAGCGCGCGCCGCGGACACTTTTCTCATTTCAAGCGCTCGGGGGCTTTAGCTACTC	1425
QY	481	AGAGCGGAGCGCCAGAGTAAACAGGAGCTTGACGCGCTTCTACTAGTCGCTAGGCTTGAGA	540
Db	1426	AGAGCGGAGCGCCAGAGTAAACAGGAGCTTGACGCGCTTCTACTAGTCGCTAGGCTTGAGA	1485
QY	541	CTTGGGAACTCCGACCGCGCGGCAACAGCGGCTGAGGGAGCGCTTGGAGAAAGCCGAGCGGCAAGC	600
Db	1486	CTTGGGAACTCCGACCGCGCGGCAACAGCGGCTGAGGGAGCGCTTGGAGAAAGCCGAGCGGCAAGC	1545
QY	601	GC CGGAGCGCGAGATCAGCGGGCGGAAAGGGGCGCGAGTTTTCGCGAGGAGGACGCTGC	660
Db	1546	GC CGGAGCGCGAGATCAGCGGGCGGAAAGGGGCGCGAGATTTTTCGCGAGGAGGACGCTGC	1605

QY 661 CGCCAGAGTGGAGAGCCGCCCTCTGTCACTTGGAGCACTGTGCGCTGCTCT 720
 Db 1606 CGCCAGAGTGGAGAGCCGCCCTCTGTCACTTGGAGCACTGTGCGCTGCTCT 1665
 QY 721 CCCCGCTGGCCCAAGTTCTGTGGCTGTGGTGGGCTGGCCAAAGGCGGCGTGGCAC 780
 Db 1666 CCCCGCTGGCCCAAGTTCTGTGGCTGTGGTGGGCTGGCCAAAGGCGGCGTGGCAC 1725
 QY 781 TGCCTTGTGCCCAACCGCCCTGCGCGCGGCGGCCCTGTGCACTGCTCGCACTCGG 840
 Db 1726 TGCCCTTGTGCCCAACCGCCCTGCGCGCGGCGGCCCTGTGCACTGCTCGCACTCGG 1785
 QY 841 CGCGCGCGGCTGTGGTGGGCGGCGGAGTTTCTGGAGTCTTGAGTCCCTGGAGCCGAGCTGCGCGC 900
 Db 1786 CGCGCGCGGCTGTGGTGGGCGGCGGAGTTTCTGGAGTCTTGAGTCCCTGGAGCCGAGCTGCGCGC 1845
 QY 901 CCTGAGAGCCATGAGGCTCCACCTGTGGGCTGCAAGGCCCAAGAACCCACCTGCTGGAAT 960
 Db 1846 CCGTGAAGCCATGAGGCTCCACCTGTGGGCTGCAAGGCCCAAGAACCCACCTGCTGGAAT 1905
 QY 961 TAGCCATTTCGTGCTGAGTGTCCGCTGAAAGTGAATGGGCAAGTCCAGGATACCTCTC 1020
 Db 1906 TAGCCATTTCGTGCTGAGTGTCCGCTGAAAGTGAATGGGCAAGTCCAGGATACCTCTC 1965
 QY 1021 TTCCGCCAGAGCATPAACAGACAGTGTGCTTACCTTCACTCTGACACAGAGGCT 1080
 Db 1966 TTCCGCCAGAGCATPAACAGACAGTGTGCTTACCTTCACTCTGACACAGAGGCT 2025
 QY 1081 CCCCAAGGCTGTGCTGAGTCACTGTGAAGATCTGCAATGCGAAGGCTTCTATCAGCT 1140
 Db 2026 CCCCAAGGCTGTGCTGAGTCACTGTGAAGATCTGCAATGCGAAGGCTTCTATCAGCT 2085
 QY 1141 GGTGTGTGTCCACCAAGAGATGTGATCTACCTGCGCCCTCCCACTTACCAACATGTCG 1200
 Db 2086 GGTGTGTGTCCACCAAGAGATGTGATCTACCTGCGCCCTCCCACTTACCAACATGTCG 2145
 QY 1201 TTCCCTGTGGGCACTGTGGGCTGTGATGGGCAATTGGGGCCACAGTGTGCTGAATCCAA 1260
 Db 2146 TTCCCTGTGGGCACTGTGGGCTGTGATGGGCAATTGGGGCCACAGTGTGCTGAATCCAA 2205
 QY 1261 GTTCTCGGCTGTGATGTTCTGGGAAGATTGCCAGACAGACAGGCTGACGCTTCCAGTA 1320
 Db 2206 GTTCTCGGCTGTGATGTTCTGGGAAGATTGCCAGACAGACAGGCTGACGCTTCCAGTA 2265
 QY 1321 CATTGGGAGCTGTGCCGATACCTGTCAACAGGCCCGGAGAGGCAAGGCAAGCTGGCCA 1380
 Db 2266 CATTGGGAGCTGTGCCGATACCTGTGTCAACAGGCCCGGAGAGGCAAGGCAAGCTGGCCA 2325
 QY 1381 TAAAGTCCGCTGCGCAGTGGGCAAGCGGCTGCGCCAGATACCTGGAGCGCTTTGTGCG 1440
 Db 2326 TAAAGTCCGCTGCGCAGTGGGCAAGCGGCTGCGCCAGATACCTGGAGCGCTTTGTGCG 2385
 QY 1441 GCGCTTGGGGCCCTGCGAGGTGTGGAGACATATGATGACATGACAGAGGCAACGTGGCAC 1500
 Db 2386 GCGCTTGGGGCCCTGCGAGGTGTGGAGACATATGATGACATGACAGAGGCAACGTGGCAC 2445
 QY 1501 CATCACTACACAGAGACAGCGGGGCGCTGTGGGGCGTCTTCTGCTTTTCAAGCATAT 1560
 Db 2446 CATCACTACACAGAGACAGCGGGGCGCTGTGGGGCGTCTTCTGCTTTTCAAGCATAT 2505
 QY 1561 CTTCCTCTCTCTCTGATTTGGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
 Db 2506 CTTCCTCTCTCTCTGATTTGGTATGATGATGATGATGATGATGATGATGATGATGATGAT 2565
 QY 1621 GGGGCACTGTATGGCCACATCTCCAGTGTAGCCAGAGGCTGTGGTGGCCCGGTATAGCA 1680
 Db 2566 GGGGCACTGTATGGCCACATCTCCAGTGTAGCCAGAGGCTGTGGTGGCCCGGTATAGCA 2624
 QY 1681 GCAAGTCCCATTTCTGGGCTATGCTGGGCGGCGAGAGTGTGCCCAAGGAGAAAGTTGCTAAA 1740
 Db 2625 GCAAGTCCCATTTCTGGGCTATGCTGGGCGGCGAGAGTGTGCCCAAGGAGAAAGTTGCTAAA 2684
 QY 1741 GGATGTCTTCCGCGCTGGGAGATGTTTCTTCAACACTGGGAGCTGTGCTGCTGAGTGA 1800

Db 2685 GGATGTCTTCCGCGCTGGGAGATGTTTCTTCAACACTGGGAGCTGTGCTGCTGAGTGA 2744
 QY 1801 CCAGGTTTCTCCGCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
 Db 2745 CCAGGTTTCTCCGCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2804
 QY 1861 TGTGGCCACACCGAGAGTGGCAGAGGCTTTTCAAGGCTTATGATTTTCTTCAAGAGGTGA 1920
 Db 2805 TGTGGCCACACCGAGAGTGGCAGAGGCTTTTCAAGGCTTATGATTTTCTTCAAGAGGTGA 2864
 QY 1921 CGTCTATGAGTCACTGTGCGAGGCGGATGAGAGGCTGGAATGAGAGGCTTATGATTTCT 1980
 Db 2865 CGTCTATGAGTCACTGTGCGAGGCGGATGAGAGGCTTGAATGAGAGGCTTATGATTTCT 2924
 QY 1981 GCGTCCCGCCCAAGCTTTTGAACCTTATGAGGCTTATGAGGCTTATGAGGCTTATGAGGCT 2040
 Db 2925 GCGTCCCGCCCAAGCTTTTGAACCTTATGAGGCTTATGAGGCTTATGAGGCTTATGAGGCT 2984
 QY 2041 ACCTTATGCGCGGCGGCGGATTCCTGAGGCTTCAAGAGGCTTATGAGGCTTATGAGGCT 2100
 Db 2985 ACCTTATGCGCGGCGGCGGATTCCTGAGGCTTCAAGAGGCTTATGAGGCTTATGAGGCT 3044
 QY 2101 CAAACAGCAAGAAAGTTGAGTGGCAATGAGGCTTCAAGGCTTATGAGGCTTATGAGGCT 2160
 Db 3045 CAAACAGCAAGAAAGTTGAGTGGCAATGAGGCTTCAAGGCTTATGAGGCTTATGAGGCT 3104
 QY 2161 ACTGTACGTTCTGAGCAAGGCTGTAGGCTCTACCTGCGCCCTTCAACACTGCGCGTACAG 2220
 Db 3105 ACTGTACGTTCTGAGCAAGGCTGTAGGCTCTACCTGCGCCCTTCAACACTGCGCGTACAG 3164
 QY 2221 CGCCCTCTGTGGCAGAGAACTTCAAGTGTGAGAACTTCAACACTGAGGCAAGGCTGAGAA 2280
 Db 3165 CGCCCTCTGTGGCAGAGAACTTCAAGTGTGAGAACTTCAACACTGAGGCAAGGCTGAGAA 3224
 QY 2281 GGAACCTGTGTGGGCGGCGGCTGTGAGGCTGTAGGCTGTGAGGCTGTGAGGCTGTGAGG 2340
 Db 3225 GGAACCTGTGTGGGCGGCGGCTGTGAGGCTGTAGGCTGTGAGGCTGTGAGGCTGTGAGG 3284
 QY 2341 CCAGAACTGCGGCTCACTATTTTGTAAATGATGATGATGATGATGATGATGATGATGATGAT 2399
 Db 3285 CCAGAACTGCGGCTCACTATTTTGTAAATGATGATGATGATGATGATGATGATGATGATGAT 3343

RESULT 10
 AX882240
 LOCUS 2314 bp DNA linear PAT 17-DBC-2003
 DEFINITION Sequence 17145 from Patent EP1074617.
 AX882240
 ACCESSION AX882240
 VERSION AX882240.1 GI:40037055
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCES
 1 Ota, T., Isegai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J., Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K., and Otsuki, T.
 Primers for synthesizing full-length cDNA and their use
 Patent: EP 1074617-A 17145 07-FEB-2001;
 Research Association for Biotechnology (JP)
 location/Qualifiers
 1. 2314
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 60..2162
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAE91498.1"
 /db_xref="GI:40037056"
 /translation="MFASGMQTVPIEAGSVAALLLLPLILLPLLLKHLMPOLR
 MPADIAFAVRAVLCCKRALRARAALAAAAADPEPGGGLAWRLAEHLAQCSAAHFLI

Qy	93	CTGCAGGCGCCAGAAACCCACCTGCTGGAATTAGCAATTGCTGGCTGAAGTGTCCGCTG	989
Db	841	CTGCAGGCGCCAGAAACCCACCTGCTGGAATTAGCAATTGCTGGCTGAAGTGTCCGCTG	900
Qy	990	AAGTGGATGGGCGAGTGCAGAGTATCCCTCTTATCCGCCGAGACATTAACGACGCTGCC	104
Db	901	AAGTGGATGGGCGAGTGCAGAGTATCCCTCTTATCCGCCGAGATTAACGACGACCTGCC	960
Qy	1050	TGTATCATTTCACTCTGACACGAGGCTCCCTCCAGAGGTGCTCGGATCAGTCATCTGA	110
Db	961	TGTATCATTTTACCTCTGACACGAGGCTCCCTCCAGAGGTGCTCGGATCAGTCATCTGA	102

Db 1021 AGATCCGCAATGCCAGGCTTCTATCAGCTGTGTGCTCCACAGGAAGATGATCT 108

Db	1081	ACCTGCGCCCTCCCACTACCAACATGTCGGGTTCCTCGCTGGGCATTCGTGGGCTGCATGTG	1144
QY	1230	GCATTGGGGGCCACAGTGTGTGCTGAATTCCAAATTCTCGGTGGTCAAGTTCCTGGGAAGATT	1281
Db	1141	GCATTGGGGGCCACAGTGTGTGCTGAATTCCAAATTCTCGGTGGTCAAGTTCCTGGGAAGATT	1201
QY	1290	GCCACGACACAGAGGTGACGGTGTTCAGTATCAATTGGGAGCTGTCCGATACCTTGTCA	1341
Db	1201	GCCACGACACAGAGGTGACGGTGTTCAGTATCAATTGGGAGCTGTCCGATACCTTGTCA	1261
QY	1350	ACCAGCCCCCGAGCAGAGGCAAGCGTGGCCATAAGTGTCGGCTGGCACTGTGGCAGCGGC	1401
Db	1261	ACCAGCCCCCGAGCAGAGGCAAGCGTGGCCATAAGTGTCGGCTGGCACTGTGGCAGCGGC	1320
QY	1410	TGGCGCCAGATACCTGGGAGCGTTTTTGTGGGGGGCTTCGGGGCCCTGCAGAGGTCTGGAGA	1461
Db	1321	TGGCGCCAGATACCTGGGAGCGTTTTTGTGGGGGGCTTCGGGGCCCTGCAGAGGTCTGGAGA	1381
QY	1470	CATTATGACCTGACAGAGGGGCAACCTGGCCACCATCAACTACACAGACAGCGGGGCGCTG	1521
Db	1381	CATTATGACCTGACAGAGGGGCAACCTGGCCACCATCAACTACACAGACAGCGGGGCGCTG	1441
QY	1530	TGGGGCGTGTCTTCGTGCTTTACAAGATATCTTCCCTCTCTCTTGATTCCGTATGATG	1581
Db	1441	TGGGGCGTGTCTTCGTGCTTTACAAGATATCTTCCCTCTCTCTTGATTCCGTATGATG	1501
QY	1590	TCACCAACAGAGAGCCAAATTCGGGAGCCCGAGGGGGCACTGTATGGGCAATCTCCAGGTG	1641
Db	1501	TCACCAACAGAGAGCCAAATTCGGGAGCCCGAGGGGGCACTGTATGGGCAATCTCCAGGTG	1561
QY	1650	AGCCAGGAGCTCTGTGTGGCCCCCGGTAAACAGACAGAGTCCCATTCCTGGGCTATGCTGGG	1701
Db	1561	AGCCAGGAGCTCTGTGTGGCCCCCGGTAAACAGACAGAGTCCCATTCCTGGGCTATGCTGGG	1621
QY	1710	GCCAGAGCTGGCCCGAGGGAGATTGCTTAAGATGTCTTCGGGCTGGGGATGTTTTCT	1761
Db	1621	GCCAGAGCTGGCCCGAGGGAGATTGCTTAAGATGTCTTCGGGCTGGGGATGTTTTCT	1681
QY	1770	TCACAACGTGGGAGCCTGTGTGTGCTGGCATGACCAAGTTTTCTCCGCTTCGATGATCGTA	1821
Db	1681	TCACAACGTGGGAGCCTGTGTGTGCTGGCATGACCAAGTTTTCTCCGCTTCGATGATCGTA	1741
QY	1830	CTGGAGACACCTTCAGGTGAGAGGGGAGAAATGTGGCAACAACGAGGTGGCAGAGTCT	1881
Db	1741	CTGGAGACACCTTCAGGTGAGAGGGGAGAAATGTGGCAACAACGAGGTGGCAGAGTCT	1801
QY	1890	TCGAGGCGCTTGAATTTTCTTCAGAGAGGTGAACCTTATGAGATCACTGTGCCAGGGCAGTG	1941
Db	1801	TCGAGGCGCTTGAATTTTCTTCAGAGAGGTGAACCTTATGAGATCACTGTGCCAGGGCAGTG	1861
QY	1950	AAGGCAAGGCTGAGATGAGACAGCCCTAGATTCTGGGTGCCCGCCACGCTTGGACCTTATGC	2001
Db	1861	AAGGCAAGGCTGAGATGAGACAGCCCTAGATTCTGGGTGCCCGCCACGCTTGGACCTTATGC	1921

Db 1921 AGCTTACACCCAGCTGTCTGAGAACTTGGACCTTATGCCCCGATTCCTCAGGC 1980
Qy 2070 TCCAGAGCTCTTTTGGCCACACAGAGACCTTCAAAACAGAGAAAGTTCCGATGGCAATG 2129
Db 1981 TCCAGAGCTCTTTTGGCCACACAGAGACCTTCAAAACAGAGAAAGTTCCGATGGCAATG 2040
Qy 2130 AGGGCTTGACCCAGACCCCTGTCTGACCCAGCTGTACCTGTGACAGAGCTGTAGTG 2189
Db 2041 AGGGCTTGACCCAGACCCCTGTCTGACCCAGCTGTACCTGTGACAGAGCTGTAGTG 2100
Qy 2190 CCTACCTGCCCCCTGACACTGCCCCGATACAGCGCCCTCTGGAGAGAACTTGGAACTCT 2249
Db 2101 CCTACCTGCCCCCTGACACTGCCCCGATACAGCGCCCTCTGGAGAGAACTTGGAACTCT 2160
Qy 2250 GAGAACTTCCACACTGAGGCACTGAGAGAGAACTGTGGGTGGGGCCCTTGCAG 2309
Db 2161 GAGAACTTCCACACTGAGGCACTGAGAGAGAACTGTGGGTGGGGCCCTTGCAG 2220
Qy 2310 GTGACTGGGCTGTGAGGAGATCTTTTCTATACAGAACTGCGTCACTATTTGTAATA 2369
Db 2221 GTGACTGGGCTGTGAGGAGATCTTTTCTATACAGAACTGCGTCACTATTTGTAATA 2280
Qy 2370 ATGTGCTGAGAGCTGATCCAGCTGTCTGTGACT 2403
Db 2281 ATGTGCTGAGAGCTGATCCAGCTGTCTGTGACT 2314

RESULT 11

LOCUS BD159613 2314 bp DNA linear PAT 17-JAN-2003
DEFINITION Primer for synthesizing full-length cDNA and use thereof.

ACCESSION BD159613
VERSION BD159613.1 GI:27865371
KEYWORDS JP 2002191363-A/14456

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 2314)
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Ota, T., Isegai, T., Nishikawa, T., Hayashi, K., Saio, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.

TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002191363-A 14456 09-JUL-2002;
HELIIX RESEARCH INSTITUTE

COMMENT

OS Homo sapiens (human)
PN JP 2002191363-A/14456
PD 09-JUL-2002
PI 28-JUL-2000 JP 2000280990
PI TOSHIO OTA, TAKAO ISEGA, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
PI SHITO,
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI KEIICHI NAGAI, TETSUJI OTSUKI

PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12N5/10, C12P21/02, C12Q1/68, C12P21/08, G06F17/30, C12N5/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers (60) . (2159) .
FT CDS Location/Qualifiers
1..2314
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 96.1%; Score 2310.8; DB 6; Length 2314;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2312; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 90 GGAAGAGAAAGTCTCAGCTAGAGAGAGCGCCCTTGGTTTTCGGAAGGAGATCAGGGA 149
|||

Db 1 GGGAGAGAAAGTCTCAGCTAGAGAGAGCGCCCTAGGTTTTCGGAAGGAGATCAGGGA 60
Qy 150 TGTTTTCGAGCGGCTGGAAACAGACGGTCCCATAGAGAAACGGGCTCCATGCGTCCC 209
Db 61 TGTTTTCGAGCGGCTGGAAACAGACGGTCCCATAGAGAAACGGGCTCCATGCGTCCC 120
Qy 210 TCTGTGCTGCCCCCTGTGCTGTGTCTTACCCCTGCTGTGCTGAAGTACACTTGTGC 269
Db 121 TCTGTGCTGCCCCCTGTGCTGTGTCTTACCCCTGCTGTGCTGAAGTACACTTGTGC 180
Qy 270 CGCAGTTGCGTGGCTTCCGCGGAACTTGGCTTTTGGGTGGAGTCTGTGTCTGAATA 329
Db 181 CGCAGTTGCGTGGCTTCCGCGGAACTTGGCTTTTGGGTGGAGTCTGTGTCTGAATA 240
Qy 330 GGGCTTTCGAGCTGCGGCTTGGCCGCGGCTGCGGCGGACCCGGAAGGTTCCGAGGGGG 389
Db 241 GGGCTTTCGAGCTGCGGCTTGGCCGCGGCTGCGGCGGACCCGGAAGGTTCCGAGGGGG 300
Qy 390 GCTGACGCTTGGCTGGCGCTTCCGCGAACTGGCCACAGCGCGCCGCGCACTCTTTC 449
Db 301 GCTGACGCTTGGCTGGCGCTTCCGCGAACTGGCCACAGCGCGCCGCGCACTCTTTC 360
Qy 450 TCAATTACAGGCTCGCGGCGCTTAACTACTCAGAGCGGAGGCGAGAGTAAACAGGCGTG 509
Db 361 TCAATTACAGGCTCGCGGCGCTTAACTACTCAGAGCGGAGGCGAGAGTAAACAGGCGTG 420
Qy 510 CACGCGCTTTCCTACCTGCTGCTAGAGCTGGGAGCTGGGAGCCGACGCGCGGACAGCGGG 569
Db 421 CACGCGCTTTCCTACCTGCTGCTAGAGCTGGGAGCTGGGAGCCGACGCGCGGACAGCGGG 480
Qy 570 AGGGAGCGCTGGAGAGAGCGGAGCGGCGGCGGAGCCGGAATGAGCGGCGGGA 629
Db 481 AGGGAGCGCTGGAGAGAGCGGAGCGGCGGAGCCGGAATGAGCGGCGGGA 540
Qy 630 GCGGCGCGAGTTTGGCCGAGAGGAGCGGTGCGGCGGAGAGTGAAGAGCGGCGGCGCTC 689
Db 541 GCGGCGCGAGTTTGGCCGAGAGGAGCGGTGCGGCGGAGAGTGAAGAGCGGCGGCGCTC 600
Qy 690 TGTACCTGAGAGCACTGTGGCGCTGCTCTCCCGGCTGGCCGAGATTCTGTGTGCTCT 749
Db 601 TGTACCTGAGAGCACTGTGGCGCTGCTCTCCCGGCTGGCCGAGATTCTGTGTGCTCT 660
Qy 750 GGTTCGCGCTGGCGAGAGCGGCGGCTGGCGAGCTTTTGGCCAGCGGCGGCGGCGG 809
Db 661 GGTTCGCGCTGGCGAGAGCGGCGGCTGGCGAGCTTTTGGCCAGCGGCGGCGGCGG 720
Qy 810 GCGGCGCTGTGCACTGCTCCGAGCTGCGGCGGCGGCGGCGGCTGTGTGTGGCGGAGCT 869
Db 721 GCGGCGCTGTGCACTGCTCCGAGCTGCGGCGGCGGCGGCGGCTGTGTGTGGCGGAGCT 780
Qy 870 TTTTGGAGTTCCTTGGAGCGGAGCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 929
Db 781 TTTTGGAGTTCCTTGGAGCGGAGCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
Qy 930 CTGAGGCGGAG 989
Db 841 CTGAGGCGGAG 900
Qy 990 AAGTGAATGGCGAGTGGCGAGATACCTCTTCCCGGAGAGATTAAGAGAGAGAGAGAG 1049
Db 901 AAGTGAATGGCGAGTGGCGAGATACCTCTTCCCGGAGAGATTAAGAGAGAGAGAGAG 960
Qy 1050 TGTACATCTTCACTCTGAG 1109
Db 961 TGTACATCTTCACTCTGAG 1020
Qy 1110 AGATCTGCAATGCAAGGCTTCTATCAGCTGTGTGTGTCAACAGAGAGATGTATCT 1169
Db 1021 AGATCTGCAATGCAAGGCTTCTATCAGCTGTGTGTGTCAACAGAGAGATGTATCT 1080
Qy 1170 ACTTCGCGCTCCAGCTTACAGATGCGGTTTCCGTTGGGCACTGGGCGGCGGCGGCGG 1229
Db 1081 ACTTCGCGCTCCAGCTTACAGATGCGGTTTCCGTTGGGCACTGGGCGGCGGCGGCGG 1140

QY 1230 GCATTGGGGCCACAGTGTGTCTGAATTCAGATTCTCGGCTGTGTCAGATTCTGGGAAAGTT 1289
 DB 1141 GGAATGGGGCCACAGTGTGTCTGAATTCAGATTCTCGGCTGTGTCAGATTCTGGGAAAGTT 1200
 QY 1290 GCAGAGAGACAGGGGAGAGGTGTTCAGATCATTTGGGAGGTGTGCCGATACCTTCA 1349
 DB 1201 GCCAGACAGACAGGGTGTGACGGTGTTCAGATCATTTGGGAGGTGTGCCGATACCTTCA 1260
 QY 1350 ACCAGCCCCGAGCAGAGGAGAGAGTGTGCGCATAGGTGCGCTGAGCATGTGGCAGCGGGC 1409
 DB 1261 ACCAGCCCCGAGCAGAGGAGAGAGTGTGCGCATAGGTGCGCTGAGCATGTGGCAGCGGGC 1340
 QY 1410 TGGGCCAGATACCTTGGGAGGCTTTTGTGCGCGGCTTGTGGGCGCTGAGAGGTGTGAGGA 1469
 DB 1321 TGGGCCAGATACCTTGGGAGGCTTTTGTGCGCGGCTTGTGGGCGCTGAGAGGTGTGAGGA 1380
 QY 1470 CATATGACTGACAGAGGGCAAGTGTGCGCAGCATCACTACACAGAGAGCGGGCGCTG 1529
 DB 1381 CATATGACTGACAGAGGGCAAGTGTGCGCAGCATCACTACACAGAGAGCGGGCGCTG 1440
 QY 1530 TGGGGGTGTCTCTGCTGCTTTCAGAGCATATCTTCCCTCTCTCTTATTTGGCTATGATG 1589
 DB 1441 TGGGGGTGTCTCTGCTGCTTTCAGAGCATATCTTCCCTCTCTCTTATTTGGCTATGATG 1500
 QY 1590 TCACCAAGAGAGAGCCAAATTCGGAGCCCCCAGGGGCACTGTATGCGCATCTCCAGGTG 1649
 DB 1501 TCACCAAGAGAGAGCCAAATTCGGAGCCCCCAGGGGCACTGTATGCGCATCTCCAGGTG 1560
 QY 1650 AAGCAGGGCTGTGTGTGTGCGCGGCTGAGCAGCAGTGTCCCATCTCTGCGCTATGTCTGCG 1709
 DB 1561 AAGCAGGGCTGTGTGTGTGCGCGGCTGAGCAGCAGTGTCCCATCTCTGCGCTATGTCTGCG 1620
 QY 1710 GGGCAGAGCTGCGCCAGGGGAAAGTTGCTAAAGATGTCTTCCGCGCTGCGGAGTGTCTTCT 1769
 DB 1621 GGGCAGAGCTGCGCCAGGGGAAAGTTGCTAAAGATGTCTTCCGCGCTGCGGAGTGTCTTCT 1680
 QY 1770 TCACAACCTGGGAGACCTGTCTGTCTGCGATGACCAAGTGTCTTCCGCTTCATGATGCTA 1829
 DB 1681 TCACAACCTGGGAGACCTGTCTGTCTGCGATGACCAAGTGTCTTCCGCTTCATGATGCTA 1740
 QY 1830 CTGGAGACACCTTCAAGTGTGAGAGGGGAGAAAGTGTGCCACAACGAGGTGTGCGAGAGGCT 1889
 DB 1741 CTGGAGACACCTTCAAGTGTGAGAGGGGAGAAAGTGTGCCACAACGAGGTGTGCGAGAGGCT 1800
 QY 1890 TCGAGGCTCTAGATTTTCTTTCAGAGAGTGAAGTCTATGAGTCACTGTGCGAGGCGCATG 1949
 DB 1801 TCGAGGCTCTAGATTTTCTTTCAGAGAGTGAAGTCTATGAGTCACTGTGCGAGGCGCATG 1860
 QY 1950 AAGGCGAGGCTGGAATGCGAGCCCTAGTCTGCGTCCCGCCCAAGCTTTCAGTCTTATGC 2009
 DB 1861 AAGGCGAGGCTGGAATGCGAGCCCTAGTCTGCGTCCCGCCCAAGCTTTCAGTCTTATGC 1920
 QY 2010 AACTCTACACCCAGCGTGTCTGAGAACTTGCACCTTATGCGCGGCCGATTCCTCAAGGC 2069
 DB 1921 AACTCTACACCCAGCGTGTCTGAGAACTTGCACCTTATGCGCGGCCGATTCCTCAAGGC 1980
 QY 2070 TCGAGAGTCTTTGGCCACACAGAGACCTTCAACAGAGAAAGTTGGAATGGAATG 2129
 DB 1981 TCGAGAGTCTTTGGCCACACAGAGACCTTCAACAGAGAAAGTTGGAATGGAATG 2040
 QY 2130 AAGGCTTGCAGCCAGACACCTGTCTGACCCACCTGTAAGTTCGGAAGGCTGTAAGT 2189
 DB 2041 AAGGCTTGCAGCCAGACACCTGTCTGACCCACCTGTAAGTTCGGAAGGCTGTAAGT 2100
 QY 2190 CCTACTGTGCTTCAACATGTGCGCGGTAAGAGCGCCCTCTGAGCAGAAACCTTGAATCT 2249
 DB 2101 CCTACTGTGCTTCAACATGTGCGCGGTAAGAGCGCCCTCTGAGCAGAAACCTTGAATCT 2160
 QY 2250 GAGAACTTCAACACTGTGAGGAGCCTGAGAGAGGAACTGTGTGGGTGTGGGGCGCTGTGAG 2309
 DB 2161 GAGAACTTCAACACTGTGAGGAGCCTGAGAGAGGAACTGTGTGGGTGTGGGGCGCTGTGAG 2220

QY 2310 GTGTACTGGGCTGTGAGGATCTTTTCTATACCAGACGCGGTCACTATTTTGTATATA 2369
 DB 2221 GTGTACTGGGCTGTGAGGATCTTTTCTATACCAGACGCGGTCACTATTTTGTATATA 2280
 QY 2370 ATGTGGCTGAGCTGATCCAGCTGTCTTGACCT 2403
 DB 2281 ATGTGGCTGAGCTGATCCAGCTGTCTTGACCT 2314
 RESULT 12
 AK027499
 LOCUS
 DEFINITION
 Homo sapiens cDNA FLJ14593 fis, clone NTBRM4002073, moderately
 similar to Mus musculus fatty acid transport protein 3 mRNA.
 ACCESSION
 AK027499
 VERSION
 AK027499.1 GI:14042218
 KEYWORDS
 Oligo capping; fis (full insert sequence).
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE
 AUTHORS
 1 Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y.,
 Nishikawa, T., Nagai, K., Sugano, S., Shiratori, A., Sudo, H.,
 Magatsuna, M., Hosoliri, T., Kaku, Y., Kodaira, H., Kondo, H.,
 Sugawara, M., Takahashi, M., Chiba, Y., Ishida, S., Murakawa, K.,
 Ono, Y., Takiguchi, S., Watanabe, S., Kimura, K., Murakami, K.,
 Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Makamatsu, A.,
 Nakamura, Y., Nagahara, K., Masuo, Y., Ninomiya, K. and Iwayanagi, T.
 NEDO human cDNA sequencing project
 TITLE
 NEDO human cDNA sequencing project
 JOURNAL
 Unpublished
 REFERENCE
 2 (bases 1 to 2314)
 Isogai, T. and Otsuki, T.
 DIRECT SUBMISSION
 Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute,
 Genomics Laboratory; 1532-3 Yana, Kiserazu Chiba 292-0812, Japan
 (E-mail: genomics@nri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing;
 Research Association for Biotechnology; cDNA library construction;
 5'- & 3'-end one pass sequencing and clone selection; Helix
 Research Institute (supported by Japan Key Technology Center etc.)
 and Department of Virology, Institute of Medical Science,
 University of Tokyo.
 FEATURES
 Location/Qualifiers
 source
 1. 2314
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="NT2RM4002073"
 /cell_line="NT2"
 /cell_type="teratocarcinoma"
 /note="cloning vector: pUC19FL3-mRNA from uninduced NT2
 neuronal precursor cells."
 60. 2162
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAB55156.1"
 /db_xref="GI:14042218"
 /translation="MFAAGNQTVPIEAGSYAALLLLPLLLPLLLKLLHLPOLR
 WLPAIDAFVRAICCKRALBARALAAAADPEPGGCSLARLAEILAOQRAATFLI
 HGRSPRSYSEARERENBARAFRLALGMDWGDGDSGSGASGERARAPGDAAG
 SAAPRAGCGAGARAGGAAPISPAATVALLPAGSEPIRLMWGKALKAALRAFAVPTL
 RKPPLHCLRSGLARALVAPFELSLERBDLPALAMGLHMAAGPGRHPRAGISLLA
 EYSAEVDGVPGLYSSPQITDTCLYITFSGTGLPKARISHLKILOCQGFYOLGV
 HOEDVLYALPLPHMSGLIGVCMGIGATVTLGPKARSAGQFWDCCOHRVYQYI
 GELCRILVQPPSKAERGLKVLAVGSGIRPTWRFVRRFPLOLVLETGYITENVA
 TINYQDGRGAVRASMLYKHPFSLIRDVITGSPIDPOGHCVAATFGEGRGLVAP
 VSOQSPFVGAGPILAOGLKLVDFRGGDVEFNTGDLVDCDQGLEFHEHEDTDTFR
 KNGENVATREYVAEVEALDPLDQNVNVGVTGPGHGRGAMALVLRPHALDIDMLYT
 HVENPYPARPRFLRLOSALTITETFOQKVRANEGFDSTLSLDPVLVDQAVAY
 LPLTARYSALLAGNLRIT"

OY		2190	CCGACCTGGCCCCCTCACAACATCGGCCGGTCAAGGCCTCTGGCAGAGAAACCTTGCAACT	2244
OY		2190	CCGACCTGGCCCCCTCACAACATCGGCCGGTCAAGGCCTCTGGCAGAGAAACCTTGCAACT	2244
Db		2101	CCTACCTGCCCCCTCACAACATCGGCCGGTCAAGGCCTCTGGCAGAGAAACCTTGCAACT	2166
OY		2250	GAGAACTTCACACCTGAGGACCCTGAAGAAGAACTGTGGGGTGGGGGCCGTTGCAG	2308
Db		2161	GAGAACCTTCACACCTGAGGACCCTGAAGAAGAACTGTGGGGTGGGGGCCGTTGCAG	2220
OY		2310	GTGTCTGGGGCGTGTAGGGATTTTTCTTAATACCAACCTGGCGGTACTATTGTGATCA	2365
Db		2221	GTGTCTGGGGCGTGTAGGGATTTTTCTTAATACCAACCTGGCGGTACTATTGTGATCA	2286
OY		2370	ATGTGGCTGAGACTGATCCAGCTGTCTCTGACCT	2403
Db		2281	ATGTGGCTGAGACTGATCCAGCTGTCTCTGACCT	2314
RESULT 13				
LOCUS		EC029792		
DEFINITION		Homo sapiens solute carrier family 27 (fatty acid transporter), member 3, mRNA (cDNA clone MGC:35483 IMAGE:5195359), complete cds.		
ACCESSION		BC029792		
VERSION		BC029792.1		
KEYWORDS		GI:20988247		
SOURCE		MGC.		
ORGANISM		Homo sapiens (human)		
REFERENCE				
AUTHORS				
		Homo sapiens		
		Euryarchaea; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
		1 (bases 1 to 2333)		
		Straussberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Dietchenko,L., Marusina,K., Farmer,A.C., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheper,T.E., Brownstein,M.J., Usdin,T.B., Toshituki,S., Carninci,P., Prange,C., Kana,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mulahy,S.U., Bosca,S.A., Moswan,P.J., McCrexman,R.J., Malek,J.A., Gunaratne,P.H., Richards,S., Wozley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Huliyar,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,R.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,U., Myers,R.M., Butlerfield,Y.S., Krzywnicki,M.I., Skalska,U., Smalls,D.E., Scheraga,A., Schein,J.E., Jones,S.J. and Merritt,M.A.		
TITLE		Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences		
JOURNAL		Proc Natl. Acad. Sci. U.S.A. 99 (26), 16939-16903 (2002)		
PUBMED		12477932		
REFERENCE		2 (bases 1 to 2333)		
AUTHORS		Straussberg,R.		
JOURNAL		Submitted (06-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Offices, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK		NIH-MGC Project URL: http://mgc.nci.nih.gov		
COMMENT		Contact: MGC help desk Email: cgabbs-remail.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNLT) DNA Sequencing By: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: http://www.hgsc.bcm.tmc.edu/cdna/ Contact: am@bcm.tmc.edu Gunataneu, P.H., Garcia, A.M., Lu, X., Huliyar, S.W., Louisged, H., Kwois,C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nataravali,		

A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL ac: <http://image.llnl.gov>

Series: IRAX Plate: 50 Row: n Column: 23.

Location/Qualifiers

1. .2333

source

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:35483 IMAGE:5195359"

/tissue_type="Brain, adult, 6 pooled whole brains"

/clone_id="NH_MGC_114"

/lab_host="DH10B"

/note="vector: pCMV-SPORT6"

1. .2333

gene

/gene="SLC27A3"

/note="synonyms: FATP3, MGC4365"

/db_xref="locusid:11000"

/db_xref="MIM:604193"

24. .2126

/codon_start=1

/product="SLC27A3 protein"

/protein_id="AAH29729.1"

/db_xref="GI:20988248"

/db_xref="locusid:11000"

/transcript="MFASGMQVPIVEAGSMALLPLPLLLPLLLKLMHPOUR
WLPDLFAVVALCKKRLARALAAADPEEGGCSLAWRLAEIAOORAAHFLPI
HGSFRFSYSEAESRNSRARAFLALGMDDPGDGSESSAGGERAAAGACGAAGA
SGAPAGDGAAGGAAAPLSPGATVALLPGEFLIMLPGFELIMFGGLAKKAAAFVPA
RGLDLHCSICGARALVAPLPELSPALPRLPAGALIMAGAPCHTAGSIDLIA
RSGELTCLPVPYGLSSPQSTITDTLTLYFTSGTGLPRAAISHKILIOCGFYLCGV
HQBVLVIALPLVHMSGSLIGVCSWIGATVLSKFSAGQEWEDQQRHVIVFYTH
GELIRVLNQPSKAEGRHKVRLAVSGLEPDETERVRRFGPLQVLETGLTEGNVAP
TINTGORGAVGRASMLYKHI FPELSLRVDTTGEPIRDPQGHMATS PGRFIDVPA
VSSQPLFVAGGPELALQGLKLDKVPFPGVFNITGDGFLRPHRTGDFPRL
WKGNNVATVAVAEFLADFLQEVNAGVAVPGHEGRAGMAALVLRPHALDMLQVAT
HVENLDPYARPRFLRQBSLATETTRKQKVRMANANGPDSLTSLDPLVLDQAVAVY
LPLTARISALMAGNLR"

579. .2015

/note="CaC; Region: Acyl-CoA synthetases
(AMP-forming)/AMP-acid ligases II [lipid metabolism /
Secondary metabolites biosynthesis, transport, and
catabolism]"

/db_xref="CDD:COG0318"

misc_feature

Query Match 94.8%; Score 2280; DB 9; Length 2333;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN

QY 126 GTTTTCGGAAGGAGCATCAGGATGTTTGGAGACGGCTGGAACCAACGCTGGCGATAG 185

Db 1 GTTTTCGGAAGGAGGATCAGGAGTGTTTGGAGACGGCTGGAACCAACGCTGGCGATAG 60

QY 186 AGAAGAGGGGCTCATGGGCTCCCTCTGTGCTGCTGCGCCCTGTGCTCTTGTGTAACGGTGC 245

Db 61 AGAAGAGGGGCTCATGGGCTCCCTCTGTGCTGCTGCGCCCTGTGCTCTTGTGTAACGGTGC 120

QY 246 TGTGCTGAAAGCTACACCTCTGCGCGGAGTTGGCGCTGGCTTCGGCGGAGCTTGGCCCTTTG 305

Db 121 TGTGCTGAAAGCTACACCTCTGCGCGGAGTTGGCGCTGGCTTCGGCGGAGCTTGGCCCTTTG 180

QY 306 CGGTCGCGAGCTCTGTGCTGCAAAAGGGCTCTTGAGCTGCGGCGCTTGGCGCGGCTGCGC 365

Db 181 CGGTCGCGAGCTCTGTGCTGCAAAAGGGCTCTTGAGCTGCGGCGCTTGGCGCGGCTGCGC 240

QY 366 CCGAAGCCGGAAGGTCCGAGAGGGGGGCTGACAGCTTGAGCTTGAGCGCTTGCGGAACTGGGCC 425

Db 241 CCGAAGCCGGAAGGTCCGAGAGGGGGGCTGACAGCTTGAGCTTGAGCGCTTGCGGAACTGGGCC 300

QY 426 AGAAGCGCGCGCGGCAACCTTTTCAATTGAAGGCTTGCGGCGCTTGAAGCTACCAAGG 485

Db	301	AGCAGCGCGCGCGACACTTTTCTCATTCACAGCTCGCGGCGGTTTAGCTACACTCAGAGG	360
OY	486	CGAGCGCGGAGATTAACAAGGGCTGCAAGCGCTTCTCTACGTGCGTTAGGGCTGGACCTGGG	545
Db	361	CGAGCGCGGAGATTAACAAGGGCTGCAAGCGCTTCTCTACGTGCGTTAGGGCTGGACCTGGG	420
OY	546	GACCAGACGGCGGCGACACAGCGCGGAGGGAGCGCTGGAGAAAGGCGAGCGGGCACAGCGCGG	605
Db	421	GAACCGAGCGCGGCGACACAGCGCGGAGGGAGCGCTGGAGAAAGGCGAGCGGGCACAGCGCGG	480
OY	606	GAGCTCGAGATGCAAGCGGCGCGAAGCGGCGCGGAGTTTGCTGGAGAGGAGCGGTGCGGCCA	665
Db	481	GAGCGCGGAGATGCAAGCGGCGCGAAGCGGCGCGGAGTTTGCTGGAGAGGAGCGGTGCGGCCA	540
OY	666	GAGGTGGAGAGAGCGCGCGCGCTCTGTATCACTGGAGGCAACTGTGGCGCTGTCTCTCCCGG	725
Db	541	GAGGTGGAGAGAGCGCGCGCGCTCTGTATCACTGGAGGCAACTGTGGCGCTGTCTCTCCCGG	600
OY	726	CTGGCGCCAGAGTTTCTGTGGCTCTGGTTCGGGACTGGCCAGACGCGGCTGTGGACACTGGCT	785
Db	601	CTGGCGCCAGAGTTTCTGTGGCTCTGGTTCGGGACTGGCCAGACGCGGCTGTGGACACTGGCT	660
OY	786	TTGTGCCCCACCGGCTCTGGCGCGGGGGCGCCCTGTGTGTCACTGGCTTCGCAAGCTGTGCGCGGC	845
Db	661	TTGTGCCCCACCGGCTCTGGCGCGGGGGCGCCCTGTGTGTGTCACTGGCTTCGCAAGCTGTGCGCGGC	720
OY	846	GCGCGCTGGTCTGGCGCGCCAGAGTTTCTGAGAGTTCCTGTGAGACCGGAGCTGTGCGCGCTGAG	905
Db	721	GCGCGCTGGTCTGGCGCGCCAGAGTTTCTGAGAGTTCCTGTGAGACCGGAGCTGTGCGCGCTGAG	780
OY	906	GAGCCATGGGGCTCCACTGTGTGGCTGTGAGGCGCCAGAACCCACCTGTCTGGAAATTAGCG	965
Db	781	GAGCCATGGGGCTCCACTGTGTGGCTGTGAGGCGCCAGAACCCACCTGTCTGGAAATTAGCG	840
OY	966	ATTGTGCTGGCTGGAAGTGTCCGCTGAAGTGGATGGGCGAGTCGACAGATACCTGTCTTCC	1025B
Db	841	ATTGTGCTGGCTGGAAGTGTCCGCTGAAGTGGAGTGGCGAGTGGCCAGAAATCCTGTCTTCC	900
OY	1026	CCCAAGACATTAACAGACACAGTGGCTGTGACATCTTTCACCTGTGGACCAACGCGGACTTCCCA	1085B
Db	901	CCCAAGACATTAACAGACACAGTGGCTGTGACATCTTTCACCTGTGGACCAACGCGGACTTCCCA	960
OY	1086	AGGCTGTCTGGATCACTGATCTGAAGATCCTTGCCATATGCCAGGGCTTCTATACGCTGTGTG	1145B
Db	961	AGGCTGTCTGGATCACTGATCTGAAGATCCTTGCCATATGCCAGGGCTTCTATACGCTGTGTG	1020
OY	1146	GTTGTCCACAGGAAGATGTGATCTTACCTGTGCGCTCCCACTGTACACATGTCCGGTCC	1205B
Db	1021	GTTGTCCACAGGAAGATGTGATCTTACCTGTGCGCTCCCACTGTACACATGTCCGGTCC	1080
OY	1206	TGCTGGGCATCGTGGGCTGTGCATGTGGCAATTTGGGGCCACAGTGTGTCTGAATTCGAATCT	1265B
Db	1081	TGCTGGGCGATCGTGGGCTGTGCATGTGGCAATTTGGGGCCACAGTGTGTCTGAATTCGAATCT	1140
OY	1266	CGGCTGTCAAGTTCGGGAAAGTTTCCACAGACACAGGGGTGACGTTTCCAGTACATTT	1325B
Db	1141	CGGCTGTCAAGTTCGGGAAAGTTTCCACAGACACAGGGGTGACGTTTCCAGTACATTT	1200
OY	1326	GGAGACTGTGCGATACCTTGTCAACCAAGCCCCGAGCAAGGACAAAGTGGCCATAAGG	1385B
Db	1201	GGAGACTGTGCGATACCTTGTCAACCAAGCCCCGAGCAAGGACAAAGTGGCCATAAGG	1260
OY	1386	TCCGGCTGTGAGTGGGGACGCGGGCTGGCCAGAACTGTGGAGAGCTTTGTGTGGCGGCT	1445B
Db	1261	TCCGGCTGTGAGTGGGGACGCGGGCTGGCCAGAACTGTGGAGAGCTTTGTGTGGCGGCT	1320
OY	1446	TCGGGCGCTGTGAGGTGTGTGAGACATATGACTACAGAGGGCAAGTGGCCACATCA	1505B
Db	1321	TCGGGCGCTGTGAGGTGTGTGAGACATATGACTACAGAGGGCAAGTGGCCACATCA	1380
OY	1506	ACTAACACAGACAGCGGGGCGGTGTGGGGCGGTCTTCTGGCTTTACAGACATATCTTC	1565B
Db	1381	ACTAACACAGACAGCGGGGCGGTGTGGGGCGGTCTTCTGGCTTTACAGACATATCTTC	1440

QY	1566	CCCTCTCCTGATTTGGCTATGATATGTCAACAACAGAGAGGCCAAATTCTGGGACCCCCAGGGGCG	1625
Db	1441	CCCTTCTCTTATTCGCTATGATATGTCAACACAGAGAGGCCAAATTCTGGGACCCCCAGGGGCG	1500
QY	1626	ACTGATATGCGCACATCTTCCAGGTAGCGCAGGGGCTGTGTGGCCCCGGGTAAAGCCAGCAGT	1685
Db	1501	ACTGATATGCGCACATCTTCCAGGTAGCGCAGGGGCTGTGTGGCCCCGGGTAAAGCCAGCAGT	1560
QY	1686	CCCCCATTTCTTGGGCTATAGCTGCGCGGGCCAGAGCTGCGGCCCAAGGGGAAGTTGCTAAAGATG	1745
Db	1561	CCCCCATTTCTTGGGCTATAGCTGCGCGGGCCAGAGCTGCGGCCCAAGGGGAAGTTGCTAAAGATG	1620
QY	1746	TCTTCGCGGCTGGGGATAGTTTCTTCAACAACCTGAGGAGCCTGCTGTCGCATATCCAAAG	1805
Db	1621	TCTTCGCGGCTGGGGATAGTTTCTTCAACAACCTGAGGAGCCTGCTGTCGCATATCCAAAG	1680
QY	1806	GTTTTCTTCGCTTCATATGATGTACTGTAAGACACCTTCAAGTGTGAAGGGGGAGATGTGG	1865
Db	1681	GTTTTCTTCGCTTCATATGATGTACTGTAAGACACCTTCAAGTGTGAAGGGGGAGATGTGG	1740
QY	1866	CCAACAACGAGAGTGGCAGAGGTTCTTCAGAGCCCTAGATTTTCTTCAGAGGTGAACGCT	1925
Db	1741	CCAACAACGAGAGTGGCAGAGGTTCTTCAGAGCCCTAGATTTTCTTCAGAGGTGAACGCT	1800
QY	1926	ATGAGATCACTGTGCCAGGCGATTAAGGCAAGGGCTGTGAATGGCAAGCCCTAATGTTCTGGCTC	1985
Db	1801	ATGAGATCACTGTGCCAGGCGATTAAGGCAAGGGCTGTGAATGGCAAGCCCTAATGTTCTGGCTC	1860
QY	1986	CCCCCGACGCTTTGACCTTATGCAAGCTCTTACACCAACGCTGTCTGAAGAATTGGCACCTT	2045
Db	1861	CCCCCGACGCTTTGACCTTATGCAAGCTCTTACACCAACGCTGTCTGAAGAATTGGCACCTT	1920
QY	2046	ATGCCCGGCCCCGATTCCTCAGGCTCCAGAGGCTCTTTGGCCACACAGAGACTTCAAAAC	2105
Db	1921	ATGCCCGGCCCCGATTCCTCAGGCTCCAGAGGCTCTTTGGCCACACAGAGACTTCAAAAC	1980
QY	2106	AGCAGAAAGTTCGAGATGAGCAATGAGGAGCTTCGACCCGACACCCCTGTCTTGACCCACTGT	2165
Db	1981	AGCAGAAAGTTCGAGATGAGCAATGAGGAGCTTCGACCCGACACCCCTGTCTTGACCCACTGT	2040
QY	2166	ACGTTCTTGAGACAGAGCTGTAGTGTGCTTACCTGCGCCCTCACAACTGCGCCGATACAGCGGCC	2225
Db	2041	ACGTTCTTGAGACAGAGCTGTAGTGTGCTTACCTGCGCCCTCACAACTGCGCCGATACAGCGGCC	2100
QY	2226	TCCTTGCGCAGAAACCTTTCGAATCTGAGAACTTCCACACTGAGGCACTCTGAGAGAGAAC	2285
Db	2101	TCCTTGCGCAGAAACCTTTCGAATCTGAGAACTTCCACACTGAGGCACTCTGAGAGAGAAC	2166
QY	2286	TCTGTGGGGTGGGGGCGGTGACAGTGTACTGGGCTGTCAAGGATCTTTTCTAATACAGA	2345
Db	2161	TCTGTGGGGTGGGGGCGGTGACAGTGTACTGGGCTGTCAAGGATCTTTTCTAATACAGA	2220
QY	2346	ACTGGGGTCACTAATTTTGTATATATAGGGCTGAGAGCGATACCAAGCTGTCTGTACCTAC	2405
Db	2221	ACTGGGGTCACTAATTTTGTATATATAGGGCTGAGAGCGATACCAAGCTGTCTGTACCTAC	2280

	RESULT 14	
AR168916		
LOCUS	AR168916	2166 bp DNA
DEFINITION	Sequence 101 from patent US 6288213.	linear
ACCESSION	AR168916	
VERSION	AR168916.1	GI:17905086
KEYWORDS		
SOURCE	unknown.	
ORGANISM	Unknown.	
REFERENCE	Unclassified.	
AUTHORS	1 (bases 1 to 2166)	
TITLE	Stahl, A., Hirsch, D. J., Lodish, H. F., Gimeno, R. E. and Tarraglin, L. A.	
JOURNAL	Fatty acid transport proteins	
FEATURES	Patent: US 6288213-A 101.11-SEP-2001;	
	Location/Qualifiers	

Db	2054	AGCGTCGTCGCTTACCTGCGCCCTCACAACTGCGCCGCTACAGCGCCCTCTTGCGAGGAA	2113
QY	2238	ACCTTCGAACTCGAAGAACTTCACACACTCGAGGCACTCGAGAGAGAACTCTGT	2290
Db	2114	ACCTTCGAACTCGAAGAACTTCACACACTCGAGGCACTCGAGAGAGAACTCTGT	2166
RESULT 15			
LOCUS	AR193703	2166 bp	DNA
DEFINITION	Sequence 101 from patent US 6348321.	linear	PAT 20-APR-2002
ACCESSION	AR193703		
VERSION	AR193703.1	GI:20240295	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 2166)		
AUTHORS	Steinl,A., Hirsch,D.J., Lodish,H.F., Gimeno,R.E. and Tartaglia,L.A.		
TITLE	Methods of identifying agents inhibiting fatty acid transport proteins		
JOURNAL	Patent: US 6348321-A 101 19-FEB-2002;		
FEATURES	location/Qualifiers		
source	1..2166		
	/organism="unknown" ..		
	/mol_type="unassigned DNA"		

Query Match	88.1%	Score 2119.4	DB 6	Length 2166
Best Local Similarity	99.7%	Pred. No. 0		
Matches 2146	Conservative	0	Mismatches 1	Indels 6
				Gaps 2
QY	144	CAGGGATTTTGCAGCGGCTGGAAACAGACGGTCCCATATGAGGAAACGGGCTTCATCG	203	
Db	14	CGGGGATTTTGCAGCGGCTGGAAACAGACGGTCCCATATGAGGAAACGGGCTTCATCG	73	
QY	204	CTGCCCCCTCTGCTGTGCGCCCTGCGCTGCTGTGTCTAC---GCTGTGTGTGTGAAGCTAC	260	
Db	74	CTGCCCCCTCTGCTGTGCGCCCTGCGCTGCTGTGTGTCTACCGCTGTGCTGTGTGAAGCTAC	133	
QY	261	ACCTCTGGCCGCGAGTGTGCTGTGCTTCCGGCGGACTTGGCTTTTGGGTGTGAGCTCTGT	320	
Db	134	ACCTCTGGCCCGACATTTGCTGTGCTTCCGGCGGACTTGGCTTTTGGGTGTGAGCTCTGT	193	
QY	321	GCTGCAAAAAGGACTTTTCAGAGTGTGCGCCCTTGGCGCGGACTTGGCGCCAGCCGGAAGGTC	380	
Db	194	GCTGCAAAAAGGACTTTTCAGAGTGTGCGCCCTTGGCGCGGACTTGGCGCCAGCCGGAAGGTC	253	
QY	381	CCGAGGGGGCTGCAGCCTTGGCTTGGCGCTTGGCGAACTTGGCCACGACGCGCTCGCGC	440	
Db	254	CCGAGGGGGGGCTGCAGCCTTGGCGCTTGGCGCGCTTGGCGAACTTGGCCACGAGCGCGCGCGC	313	
QY	441	ACACCTTCTCATTTACAGGCTGTGCGGCGCTTTAGCTACTACAGCGGAGCGGAGATA	500	
Db	314	ACACCTTCTCATTTACAGGCTGTGCGGCGCTTTAGCTACTACAGGCGGAGCGGAGATA	373	
QY	501	ACAGGGCTGCAGCGGCTTCTCTACGCTGGGCTAGGCTGGGATCTGGGGAACCGACGGCGGGG	560	
Db	374	ACAGGGCTGCAGCGGCTTCTCTACGCTGGGCTAGGCTGGGGAATCTGGGGAACCGACGGCGGGG	433	
QY	561	ACAGCGGCGAGGAGCGCTGGAGAAAGCGACGCGGACGCGCCGGAGCCCGGAGATGCG	620	
Db	434	ACAGCGGCGAGGAGCGCTGGAGAAAGCGAGCGGCGACGCGCCGGAGCCCGGAGATGCG	493	
QY	621	CGGCGCGAAGGGGCGCGGAGTTTGGCGGAGGGGAGCGGTGGCGCGAGAGGTGGAGAG---	677	
Db	494	CGGCGCGAAGGGGCGCGGAGTTTGGCGGAGGGGAGCGGTGGCGCGAGAGGTGGAGAGAGC	553	
QY	678	CGCGCGCCTCTGTGCTACCTGGAGCACTGTGGCGCTGTCTCTCCCGGTGGCCCAAGT	737	
Db	554	CGCGCGCCTCTGTGCTACCTGGAGCACTGTGGCGCTGTCTCTCCCGGTGGCCCAAGT	613	

QY	738	TTCTGTGGCTCTGGATTTGGGCTGGGCTGAGCCAGGCGGGCTGGCGATCTGCTTTGTGGCCACCG	797
Db	614	TTCTGTGGCTCTGGATTTGGGCTGGGCTGAGCCAGGCGGGCTGGCGATCTGCTTTGTGGCCACCG	673
QY	798	CCCTGGCGCCGGGGCGCCCTGTGCTGACCTGCTCCGACCTGGCGCGCGCGCTGTGTGC	857
Db	674	CCCTGGCGCGGGGGCCCCCTGTGCTGACCTGCGACCTGGCGCGCGCGCTGTGTGC	733
QY	858	TGGGCGCCAGATTTCTGGAATCCTTGGAGCTGGACCTTGCCCGCCCTGAGAGCATTGGGGC	917
Db	734	TGGGCGCCAGATTTCTGGAATCCTTGGAGCTGGAGCCTTGCCCGCCCTGAGAGCATTGGGGC	793
QY	918	TTCAACCTGTGGGCTGACAGGCCCAAGAAACCCCTGCTGGAATTAGCATTTGCTGGCTG	977
Db	794	TTCAACCTGTGGGCTGACAGGCCCAAGAAACCCCTGCTGGAATTAGCATTTGCTGGCTG	853
QY	978	AAATGTCCGCTGAAGTGGATGGGCGAGTGGCCAGATTACTCTCTTCCGCCAGAGCATAA	1037
Db	854	AAATGTCCGCTGAAGTGGATGGGCGAGTGGCCAGATTACTCTCTTCCGCCAGAGCATAA	913
QY	1038	CAGACAGGTGCTGTATACATCTTCACTCTGGGCAACAAGGGGCTCCCGCAAGGCTGTCGGA	1097
Db	914	CAGACAGGTGCTGTATACATCTTCACTCTGGGCAACAAGGGGCTCCCGCAAGGCTGTCGGA	973
QY	1098	TCAGTCACTGTAAGATCCTGCAATGCCAGGGCTTTCTATCAAGCTGTGTGTGTCCACAGG	1157
Db	974	TCAGTCACTGTAAGATCCTGCAATGCCAGGGCTTTCTATCAAGCTGTGTGTGTCCACAGG	1033
QY	1158	AAAGATGTATTTACTTGCCCTCCCACTCTACACATGTCCGGTTCCCTGTGGGACATCG	1217
Db	1034	AAAGATGTATTTACTTGCCCTCCCACTCTACACATGTCCGGTTCCCTGTGGGACATCG	1093
QY	1218	TGGGCTGCATGGGCAATTGGGGCCACAGTGTGTGTAATCCAGTTCTCGGCTGTCAGT	1277
Db	1094	TGGGCTGCATGGGCAATTGGGGCCACAGTGTGTGTAATCCAGTTCTCGGCTGTCAGT	1153
QY	1278	TTCTGGAAAGATTGCCAGACGACAGGGTGACGTTGTTCAATTAATTGGGAGCTGTGCC	1337
Db	1154	TTCTGGAAAGATTGCCAGACGACAGGGTGACGTTGTTCAATTAATTGGGAGCTGTGCC	1213
QY	1338	GATACCTTTGTAAACAAGCCCCCGAGCAAGCAGAAACGTGGCCATTAAGGTCGGGCTGCAG	1397
Db	1214	GATACCTTTGTAAACAAGCCCCCGAGCAAGCAGAAACGTGGCCATTAAGGTCGGGCTGCAG	1273
QY	1398	TGGGAGCGGGCTGGCCCCAGATTACTTGGAGCGTTTGTGGCGCGCTTGGGGCCCCGTC	1457
Db	1274	TGGGAGCGGGGCTGGCCCCAGATTACTTGGAGCGTTTGTGGCGCGCTTGGGGCCCCGTC	1333
QY	1458	AGGTGTCTGGAGCATATGGAATTGACAGAGGGCAACGTGGCCACCATCAATNACAAGGAC	1517
Db	1334	AGGTGTCTGGAGCATATGGAATTGACAGAGGGCAACGTGGCCACCATCAATNACAAGGAC	1393
QY	1518	AGCGGGGCGCTGTGGGGCGTGCCTTCTGTGCTTTCAAGCATATCTTCCCTCTCTCTTGA	1577
Db	1394	AGCGGGGCGCTGTGGGGCGTGCCTTCTGTGCTTTCAAGCATATCTTCCCTCTCTCTTGA	1453
QY	1578	TTCCGTATGATGTCAACAAGAGAGGCCAATTGGGACCCCCAGGGGCACTGTATGAGCA	1637
Db	1454	TTCCGTATGATGTCAACAAGAGAGGCCAATTGGGACCCCCAGGGGCACTGTATGAGCA	1513
QY	1638	CATCTCCAGGTGAGCAGGCGCTGCTGTGTGGCCCCGGTTAAGCAGAGATGCCCATTCCTG	1697
Db	1514	CATCTCCAGGTGAGCAGGCGCTGCTGTGTGGCCCCGGTTAAGCAGAGATGCCCATTCCTG	1573
QY	1698	GCTATGCTGGCGGGCCAGAGCTGGGCCCAAGGGAAATTGCTAAAGATATCTTCCGGCTG	1757
Db	1574	GCTATGCTGGCGGGCCAGAGCTGGGCCCAAGGGAAATTGCTAAAGATATCTTCCGGCTG	1633
QY	1758	GGGATGTTTTCTTCAACACTGGGACCTGTGTGCTGTGGAGTCCCAAGGTTTTCTCCGCT	1817
Db	1634	GGGATGTTTTCTTCAACACTGGGACCTGTGTGCTGTGGAGTCCCAAGGTTTTCTCCGCT	1693
QY	1818	TCCATGATCTGACTGGAGACATTTCAAGTGTGAAGGGGGAATGTGGCCACAACCGAGG	1877

```
Db 1694 TCCATATGCTACTGAGACACCTTCAGGTGGAAGGGGAGAAATGTGGCCACACCGAGG 1753
QY 1878 TGGCAGAGGTCTTCGAGGCCCTAGATTTTCTT CAGAGGTGAACGTCTATGAGTCACTG 1937
Db 1754 TGGCAGAGGTCTTCGAGGCCCTAGATTTTCTT CAGAGGTGAACGTCTATGAGTCACTG 1813
QY 1938 TGGCAGAGGTGAAGGCAAGGCTGGAATGAGAGCCCTAGTTCTGGTCCCCCCCAAGCTT 1997
Db 1814 TGGCAGAGGTGAAGGCAAGGCTGGAATGAGAGCCCTAGTTCTGGTCCCCCCCAAGCTT 1873
QY 1998 TGGACCTTATGAGCTCTACACCCAAGTGTGAGAACTTGGCACCTTATGCCCCGCCCC 2057
Db 1874 TGGACCTTATGAGCTCTACACCCAAGTGTGAGAACTTGGCACCTTATGCCCCGCCCC 1933
QY 2058 GATTCTCTAGGCTCCGAGAGTCTTTGGCCACCAAGAGACCTTCAAAACAGAGAAAGTTC 2117
Db 1934 GATTCTCTAGGCTCCGAGAGTCTTTGGCCACCAAGAGACCTTCAAAACAGAGAAAGTTC 1993
QY 2118 GATGSCAAATGAGGGCTTCGACCCAGACCCCTGTCTGACCCACTGTACGTTCTGAGCC 2177
Db 1994 GATGSCAAATGAGGGCTTCGACCCAGACCCCTGTCTGACCCACTGTACGTTCTGAGCC 2053
QY 2178 AGGCTTAGGTGCTACTGCTGCCCTTCACAATGCTCCCGGTACAGCGCCCTCTGGCAGAA 2237
Db 2054 AGGCTTAGGTGCTACTGCTGCCCTTCACAATGCTCCCGGTACAGCGCCCTCTGGCAGAA 2113
QY 2238 ACCTTGAACTTGAGAACTTCACAACCTGAGGACCTGAGAGAGGAAGTCTGT 2290
Db 2114 ACCTTGAACTTGAGAACTTCACAACCTGAGGACCTGAGAGAGGAAGTCTGT 2166
```

Search completed: July 3, 2004, 02:03:09
Job time : 9307 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 2, 2004, 22:59:20 ; Search time 6104 Seconds
(without alignment)

11765.817 Million cell updates/sec

Title: US-10-030-226-1

Sequence: 1 gcaactctcccggttctg.....tccagctctctgacttac 2405

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_pro:*
26: em_gss_phg:*
27: em_gss_vit:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1335	55.5	1990	11	AK076014 Mus muscu
2	1308	54.4	2018	11	AK037042 Mus muscu
3	942.2	39.2	997	13	EX354148 BX354148
4	939.2	39.1	1201	13	EX460000 BX460000

Result No.	Score	Query Match	Length	DB ID	Description
5	917.6	38.2	1080	13	BX358662 BX358662
6	863	35.9	930	13	AL521329 AL521329
7	861.4	35.8	994	13	BX354149 BX354149
8	855	35.6	1200	13	BX353771 BX353771
9	839.6	34.9	908	13	BQ680322 BQ680322
10	837	34.8	870	9	AL529216 AL529216
11	833.6	34.7	934	9	AL521330 AL521330
12	786.8	32.7	903	13	BQ678639 BQ678639
13	774.8	32.2	873	9	AL529217 AL529217
14	767.2	31.9	1044	12	B1820310 B1820310
15	753.6	31.3	969	13	BX419446 BX419446
16	753.4	31.3	1067	12	EM559002 EM559002
17	734.6	30.5	1298	9	AV762219 AV762219
18	733.2	30.5	775	12	B1821631 B1821631
19	729.4	30.3	1178	13	BX353772 BX353772
20	728	30.3	827	13	B1915593 B1915593
21	721.8	30.0	911	13	B1915073 B1915073
22	717.4	29.8	837	12	BG720139 BG720139
23	705.2	29.3	975	13	BQ678464 BQ678464
24	682.4	28.4	711	12	BG719288 BG719288
25	678	28.2	679	12	BM727185 BM727185
26	673.6	28.0	895	13	BQ437413 BQ437413
27	664.2	27.6	928	13	BX412740 BX412740
28	658.4	27.4	992	13	B1916319 B1916319
29	658	27.4	790	12	B1819939 B1819939
30	645.4	26.8	811	10	BE798017 BE798017
31	645.4	26.8	1216	12	BM475414 BM475414
32	642.2	26.7	671	12	B1522488 B1522488
33	641.4	26.7	684	14	CD364420 CD364420
34	634.6	26.4	1105	12	BM451631 BM451631
35	629	26.2	669	12	BM678297 BM678297
36	625.8	26.0	1143	10	BE68280 BE68280
37	619.2	25.7	676	10	BE311800 BE311800
38	619	25.7	785	10	BE277176 BE277176
39	618.6	25.7	665	12	BM681655 BM681655
40	616.6	25.6	1040	10	BE543175 BE543175
41	616	25.6	663	13	B1625560 B1625560
42	615.4	25.6	905	14	CB194469 CB194469
43	610	25.4	796	10	BE274647 BE274647
44	607.4	25.3	828	12	B1910210 B1910210
45	598.8	24.9	905	13	B1919047 B1919047

ALIGNMENTS

RESULT 1
AK076014
LOCUS
DEFINITION
MUS MUSCULUS 10 days embryo whole body cDNA, RIKEN full-length
enriched library, clone:2610030019 product:solute carrier family 27
(fatty acid transporter), member 3, full insert sequence.

ACCESSION
AK076014
VERSION
AK076014.1 GI:26344941
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
REFERENCE
1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)

REFERENCE
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Komuro, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)

TITLE
JOURNAL
MEDLINE
PUBMED
11042159

REFERENCE
AUTHORS
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komono, H., Akiyama, J., Nishi, K., Katsunuma, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujitake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN Integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL
MEDLINE
20530913
11076861

REFERENCE
AUTHORS
4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

JOURNAL
AUTHORS
5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

JOURNAL
AUTHORS
6 (bases 1 to 1990)
Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoaka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Komono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamoto, T., Yamahata, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.
Location/Qualifiers

FEATURES
source
1.1990
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:2610030U19"
/db_xref="MGI:1908071"
/db_xref="taxon:10090"
/clone="2610030U19"
/tissue_type="whole body"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 days embryo"
533..1873
/note="unlabeled protein product; putative solute carrier family 27 (fatty acid transporter), member 3 (MCD|MGI:1347358, GB|AF072758, evidence: BLASTN, 99%, match=1387)";
/codon_start=1
/protein_id="BAC36120.1"
/db_xref="GI:26344942"

ORIGIN
polya_signal
polya_site
/note="putative"
1990
/note="putative"

Query Match 55.5%; Score 1335; DB 11; Length 1990;
Best Local Similarity 82.0%; Pred. No. 5,7e-217;
Matches 1583; Conservative 0; Mismatches 300; Indels 48; Gaps 2;

331 GGCCTTCGAGCTCGAGCCCTGAGCCGAGGCTGAGCCGAGGAGGAGTCCGAGGGG 390
1 GGCCCTTCGAGCAGCGCGCTGAGCCGAGGCTGAGCCGAGGAGGAGTCCGAGGGG 60
391 CTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 450
61 GTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 120
451 CATTCAGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 510
121 CATTCAGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 180
511 AGCGCTTCCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 570
181 TCCGCTTCCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 234
571 GGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 630
235 GGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 292
631 CGCGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 690
293 -----GGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 312
691 GTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 750
313 GGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 372
751 GTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 810
373 GTTGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 432
811 CCCCTGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 870
433 ACCCTGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 492
871 TCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 930
493 CTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 552
931 TCGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 990
553 GAGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 612
991 AGTGAATGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1050
613 AGTGAATGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 672
1051 GTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1110
673 GTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 732
1111 GATCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 1170
733 GATCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 792

QY 1171 CCTGGCCCTCCCACTCTACACAGATGTCGGGTTCCCTGCTGGGCATCGTGGGCTGCATGGG 1230
 Db 793 CTTGCACTCCCACTGTAACACATGTCGGCTCCCTTCTGGGGCAATGTGGGCTGTTGGG 852
 QY 1231 CATTGGGGGCGACAGTGTGTGTAATTCAGTTCCTGGCTGTGTGTAATTCAGTTCCTGGGAAAGATG 1230
 Db 853 CATTGGGGGCGACAGTGTGTGTAATTCAGTTCCTGGCTGTGTGTAATTCAGTTCCTGGGAAAGATG 912
 QY 1291 CCAGCAGACAGAGGTGAGCGTGTTCAGTACATTTGGGAGCTGTGCCATCTTGTCAA 1350
 Db 943 CCAGAAACACAGAGGTGAGCGTGTTCAGTACATTTGGGAGTGTGCCATCTTGTCAA 972
 QY 1351 CCAGCCCCCGAGCAGAGGAGGAGAGCTGTGCTTAAGTTCGGGCTGGGAGTGGGCGAGGGCT 1410
 Db 973 CCAGCCCCCGAGCAGAGGAGGAGAGCTGTGCTTAAGTTCGGGCTGGGAGTGGGCGAGGGCT 1032
 QY 1411 GCGCCCAATACCTGGAGAGCGTTTGTGCGGCGCTTCGGGCGCCCTGCGAGCTGTGAGAC 1470
 Db 1093 GCGCCCAATACCTGGAGAGCGTTTGTGCGGCGCTTCGGGCGCCCTGCGAGCTGTGAGAC 1092
 QY 1471 ATATGACTGACAGAGGAGGAGAGCTGTGCGCAGCATCACTACAGAGACGCGGGCGCTGT 1530
 Db 1093 GTATGGCATGACAGAGGAGGAGAGCTGTGCGCAGCATCACTACAGAGACGCGGGCTGT 1152
 QY 1531 GGGGGGTGCTTCTGGCTTTCAGAGCATATCTCCCTCTCTCTGATTCGCTATGATGT 1590
 Db 1153 GGGGGGTGCTTCTGGCTTTCAGAGCATATCTCCCTCTCTCTGATTCGCTATGATGT 1212
 QY 1591 CACCAGAGAGAGCCCAATTCGGGAGCCCCAGGGGAGCTGTATGAGCAGCATTCAGAGGTGA 1650
 Db 1213 CATGACAGGGAGGCTTATTCGGAATGCCAGGGGAGCTGTATGAGCAGCATTCAGAGGTGA 1232
 QY 1651 GCCAGGSGTGTGCTGTGTCGCGCCGCTTAAGCAGCAGTCCCATCTCTGGGCTATGCTGGCGG 1710
 Db 1273 GCCAGGSGTGTGCTGTGTCGCGCCGCTTAAGCAGCAGTCCCATCTCTGGGCTATGCTGGCGG 1332
 QY 1711 GCCAGAGTGTGCGCCGAGGAGGAGTGTCTTAAGGATGTCTTCGGGCTGGGAGTGTCTTCTT 1770
 Db 1333 TCCGAGAGTGTGCGCCGAGGAGGAGTGTCTTAAGGATGTCTTCGGGCTGGGAGTGTCTTCTT 1392
 QY 1771 CAACACTGGGAGCTGTGCTGTGTCGATGACCAAGTGTCTTCGGCTTCATGATCTGTAC 1830
 Db 1393 CAATCTGGGAGCTGTGCTGTGTCGATGACCAAGTGTCTTCATCTTCAGATCTGTAC 1452
 QY 1831 TGGAGACACCTTCAAGGTGGAAGGGGAGAGATGTGGGCCACACCGAGGTGGCAGAGGTCTT 1890
 Db 1453 TGGAGACACCTTCAAGGTGGAAGGGGAGAGATGTGGGCCACACCGAGGTGGCAGAGGTCTT 1512
 QY 1891 CGAGGCGCTTGAATTTCTTCAAGAGGTGAGACGCTTATGAGTCACTGTGCCAGGCGATGA 1950
 Db 1513 GGAGACCTCTGACCTTCTTCAAGAGGTGAGACGCTTATGAGTCACTGTGCCAGGCGATGA 1572
 QY 1951 AGGCAAGGCTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2010
 Db 1573 AGGCAAGGCTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1632
 QY 2011 GCTCTACACCAAGTGTGAGAACTTTCACACTTATGCCGCGGCGCCGATCTCTCAGAGCT 2070
 Db 1633 GCTCTACACCAAGTGTGAGAACTTTCACACTTATGCCGCGGCGCCGATCTCTCAGAGCT 1692
 QY 2071 CCAGAGAGCTTGTGGCCACACAGAGAGCTTAAACAGAGAGAAAGTGTGGATGSCAAATGA 2130
 Db 1693 CCAGAGAGCTTGTGGCCACACAGAGAGCTTAAACAGAGAGAAAGTGTGGATGSCAAATGA 1752
 QY 2131 GGGCTTCAAGCCAGACACCTGTCTGAGCCCACTGACGTTTGTGAGCAGAGGCTGTAGAGTGC 2190
 Db 1753 GGGCTTCAAGCCAGACACCTGTCTGAGCCCACTGACGTTTGTGAGCAGAGGCTGTAGAGTGC 1812
 QY 2191 CTACTCTGCTTCAACATGTCGCGGCTGACAGGCGCTCTCTGAGAGAAACCTTGTGAATCTG 2250
 Db 1813 CTACTCTGCTTCAACATGTCGCGGCTGACAGGCGCTCTCTGAGAGAAACCTTGTGAATCTG 1872

QY 2251 AGAATTCCAC 2261
 Db 1873 AAACCTTCAC 1883

RESULT 2
 AK037042
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS
 TITLES
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

1 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunari, T., Taahito, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, Y., Ohara, E., Matsuda, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system - 384-format sequencing pipeline with 384 multichannel sequencer Genome Res. 10 (11), 1757-1771 (2000)

4 The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)

5 The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)

6 (bases 1 to 2018)

JOURNAL
 REFERENCE
 AUTHORS

TITLE
 AUTHORS

JOURNAL
 REFERENCE
 AUTHORS

TITLE
 AUTHORS

JOURNAL
 REFERENCE
 AUTHORS

TITLE
 JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC),

COMMENT

RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@isc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site for further details.

FEATURES

source

1. 2018

/organism="Mus musculus"

/mol_type="RNA"

/strain="C57BL/6J"

/db_xref="FANTOM:DB:9930102C17"

/db_xref="MGI:2401515"

/db_xref="taxon:10090"

/clone="9930102C17"

/sex="female"

/tissue_type="vaginal"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="adult"

<1.1901

/note="putative

solute carrier family 27 (fatty acid transporter), member

3 (MGI:MGI:1347358, GB|AF072758, evidence: BLASTN, 99%,

match=1387)"

2001.2006

/note="putative"

2018

/note="putative"

ORIGIN

Query Match 54.4%; Score 1308; DB 11; Length 2018;
 Best Local Similarity 81.1%; Pred. No. 2.3e-212;
 Matches 1588; Conservative 0; Mismatches 300; Indels 70; Gaps 3;

QY 326 AAAAGGGGCTCTTGAGAGCTCGCGCCCTGCGCGGCTGCGCGAGCCCGGAAGTCCCGAG 385
 DB 2 AAAAGGGGCTCTTGAGAGCTCGCGCCCTGCGCGGCTGCGCGAGCCCGGAAGTCTGAG 61
 QY 386 GGGGCTGAGAGCTGCGGCTGCGGCTGCGGAGTGGCCAGAGCGCGCGGAGCACC 445
 DB 62 AGGGGCTGAGAGCTGCGGCTGCGGCTGCGGAGTGGCCAGAGCGCGGAGCACC 121
 QY 446 TTCTCATTCACGAGCTGCGGCTGCTTACCTACAGAGCGGAGCGGAGCTAAACAG 505
 DB 122 TTCTCATTCACGAGCTGCGGCTGCTTACCTACAGAGCGGAGCGGAGCTAAACAG 181
 QY 506 GCTGCAAGCGCTTCTCTACGTCGCTAGGCTGAGGAGCCCGAGCGCGGAGCACC 565
 DB 182 ATTGCTCGGCGCTTCTCTACGTCGCTAGGCTGAGGAGCGCGGAGCACC 235
 QY 566 GGGGAGGGAGAGCTGAGAGAGGCGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 625
 DB 236 GAGAGGGAGAGCTGAGAGAGGCGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 295
 QY 626 GGAAGCGCGGAGGTTTCCCGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 685
 DB 296 AGA-----GGGAGCAGCCCGCGCC 313
 QY 686 CCTCTGTACCTGAGAGCACTGTGGGCTGCTCTCCCGGCTGGCCAGAGTTTCTGAG 745
 DB 314 CCTCTGTACCTGAGAGCACTGTGGGCTGCTCTCCCGGCTGGCCAGAGTTTCTGAG 373
 QY 746 CTCTGCTGGGCTGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 805
 DB 374 ATTGCTGTGAGAGTGGAGCAAGAGTGGCTGAGAGCGGAGCGGAGCGGAGCGGAGCGG 433
 QY 806 CGGGGCCCTGCTGCTGCTCCGAGCTGAGGCGCGCGCGCTGCTGCTGAGCGCC- 864

DB 434 CGAGAGACCCCTGTGTGATGCTCCCGAGCTCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 493
 QY 865 -----AGAGTTCTGAGTCCCTGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 903
 DB 494 GCTGGGTTGCTTACCTCTGTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 553
 QY 904 GAGAGCATTGGGCTCACCTGTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 963
 DB 554 GAGAGCATTGGGCTCACCTGTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 613
 QY 964 CGATTGCTGCTGAGAGTGTCCGCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1023
 DB 614 CAATTGCTATCGAGAGCAGACCAAGTGTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 673
 QY 1024 CCCCAGAGATTAACAGACAGCGCTGTATCTTCACTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1083
 DB 674 CCCCAGAGATTAACAGACAGCGCTGTATCTTCACTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 733
 QY 1084 CAGAGCTGCTGAGATGATGATGATGAGATCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1143
 DB 734 CAGAGCTGCTGAGATGATGATGATGAGATCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 793
 QY 1144 TGGTGTCCACAGAGAGATGATGATGATGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1203
 DB 794 TGGTGTCCACAGAGAGATGATGATGATGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 853
 QY 1204 CCTGTGAGGATGATGATGATGATGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1263
 DB 854 CTTTCTGAGGATGATGATGATGATGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 913
 QY 1264 CTGCGCTGCTGATGATGATGATGATGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1323
 DB 914 CTGAGCTGAGTCTGATGATGATGATGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 973
 QY 1324 TGGGAGCTGTGCGATGATGATGATGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1383
 DB 974 TGGGAGCTGTGCGATGATGATGATGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1033
 QY 1384 GATTCGCTGCGATGATGATGATGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1443
 DB 1034 GATTCGCTGCGATGATGATGATGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1093
 QY 1444 CTTGCGGCTGCTGATGATGATGATGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1503
 DB 1094 ATTGGAATCTGATGATGATGATGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1153
 QY 1504 CAATCAACAGAGCAGCGGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1563
 DB 1154 CAATCAACAGAGCAGCGGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1213
 QY 1564 CCGCTTCTGCTGATGATGATGATGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1623
 DB 1214 CCGCTTCTGCTGATGATGATGATGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1273
 QY 1624 GCACTGTATGAGCAATCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1683
 DB 1274 GCACTGTATGAGCAATCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1333
 QY 1684 GTCCCATTTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1743
 DB 1334 GTCCCATTTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1393
 QY 1744 TGTCTTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1803
 DB 1394 TGTCTTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1453
 QY 1804 AGGTTTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1863
 DB 1454 AGGTTTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1513
 QY 1864 GGCACAAACGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1923

Db 1514 GGCCAACTGAAGTGTGAGTCTTGAGAGACCTTGACCTTTCAGAGAGTGAACAT 1573
QY 1924 CTATGAGTCACTGTCGACGAGGATGAAGGAGGCTGGAAATGAGAGCCATGTCGCG 1983
Db 1574 CTATGAGTCACTGTCGACGAGGATGAAGGAGGCTGGAAATGAGAGCCATGTCGCG 1633
QY 1984 TCCCCCGCAGCTTTGAGACCTTATGACGCTTACACCAAGTGTGAGAACTTGCCACC 2043
Db 1634 GCGCCCGCAGGCTTTGAGACCTTATGACGCTTACACCAAGTGTGAGAACTTGCCACC 1693
QY 2044 TTATGCGCGCGCCGATTCCTGAGGCTCCAGAGTCTTGGCAGCAGAGACCTTCAA 2103
Db 1694 GTATCCCGACCTCGATTTCTGAGGCTCCAGAGTCTTGGCAGCAGAGACCTTCAA 1753
QY 2104 ACAGAGAAAGTTCGATGAGCAATGAGGCTTCAGACCCAGACCTGTGACCCACT 2163
Db 1754 ACAGAGAAAGTTCGATGAGCAATGAGGCTTCAGACCCAGACCTGTGACCCACT 1813
QY 2164 GTACGTTCTGAGCAAGGCTGTAGTGCTTACCTGCCCTCACACACTGCGCGTACAGCGC 2223
Db 1814 CTATGTTCTGAGCAAGATATAGGGGCTTACCTGCCCTCACACACTGCGCGTACAGTGC 1873
QY 2224 CCTCTGCGAGGAACCTTCGATCTGAGAACTTCCAC 2261
Db 1874 CCTCTGTGTGAGACCTTCGATCTGAGAACTTCCAC 1911

RESULT 3
LOCUS BX354148 997 bp mRNA linear EST 05-MAY-2003
DEFINITION BX354148 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
CDNA clone CS0DC015H21 3-PRIME, mRNA sequence.
ACCESSION BX354148
VERSION BX354148.1 GI:30373826
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 997)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 661.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgl-bin/cluster.cgi?seq=CS0DC015CD11NP1&cluster=661.f. Contact :
Feng Liang Email: fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paraday Avenue Genoscope sequence ID : CS0DC015CD11NP1.
Location/Qualifiers
source 1.
1. 997
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC015H21"
/issue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

Query Match 39.2%; Score 942.2; DB 13; Length 997;
Best Local Similarity 98.4%; Pred. No. 3.3e-150;
Matches 977; Conservative 6; Mismatches 6; Indels 4; Gaps 3;

QY 1383 AGTCCGGCTGAGAGTGGGAGGCGGCTGCGCCCAAGTACCTGGAGAGCGCTTTGTGCGC 1442
Db 990 AAGTCCGGCTGAGAGTGGG--CAGCGCGGCCCAAGTACCTGGAGAGCTTTGTGCGC 993
QY 1443 GCTTCCGGGCGCTTGAAGTGTGAGGATATGAGATATGAGGAGGAGGAGGAGGAGGAGG 1502
Db 932 GCTTCCGGGCGCTTGAAGTGTGAGGATATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 873
QY 1503 TCAACTACAGAGAGAGCGGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1562
Db 872 TCAACTACAGAGAGAGCGGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 813
QY 1563 TCCCTTCTCTGATTCGCTATGATGACCAAGAGAGGCAATTCGGAGCCCGCAGG 1622
Db 812 TCCCTTCTCTGATTCGCTATGATGATGATGATGATGATGATGATGATGATGATGATG 753
QY 1623 GGCACGTATAGGAGCAATCTCCAGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1682
Db 752 GGCACGTATAGGAGCAATCTCCAGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 693
QY 1683 AGTCCCATTCCTGAGGCTATGCTGCGGCGCAGAGCTGCGCCAGGAGGAGGAGGAGG 1742
Db 692 AGTCCCATTCCTGAGGCTATGCTGCGC--GGCAGAGCTGCGCCAGGAGGAGGAGGAGG 634
QY 1743 ATGCTTTCGCGGCTGAGGAGTGTTCCTTCAACCTGAGGAGCTGTGCTGTGATGACC 1802
Db 633 ATGCTTTCGCGGCTGAGGAGTGTTCCTTCAACCTGAGGAGCTGTGCTGTGATGACC 574
QY 1803 AAGGT-TTTCGCGCTTCATGATGCTGAGAGACACTTCAGAGGAGGAGGAGGAGGAGG 1861
Db 573 AAGGTATTTCTCGCTTCATGATGCTGAGAGACACTTCAGAGGAGGAGGAGGAGGAGG 514
QY 1862 GTGAGCAACAACGAGGAGTGGAGAGGCTTTCGAGAGGAGGAGGAGGAGGAGGAGGAGG 1921
Db 513 GTGAGCAACAACGAGGAGTGGAGAGGCTTTCGAGAGGAGGAGGAGGAGGAGGAGGAGG 454
QY 1922 GTCTATGAGATCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1981
Db 453 GTCTATGAGATCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 394
QY 1982 CGTCCCGCCAGCCTTTGAGCCTTATGAGCCTTACACCCAGCTGTCTGAGAGACTTGCGA 2041
Db 393 CGTCCCGCCAGCCTTTGAGCCTTATGAGCCTTACACCCAGCTGTCTGAGAGACTTGCGA 334
QY 2042 CCTTATGCGCGCGCGCGATTCCTCAGGCTTCAGAGGAGGAGGAGGAGGAGGAGGAGG 2101
Db 333 CCTTATGCGCGCGCGCGATTCCTCAGGCTTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 274
QY 2102 AAACAGAGAAAGTTCGAGTGGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2161
Db 273 AAACAGAGAAAGTTCGAGTGGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 214
QY 2162 CTGTAGCTTGTGAGCAGGAGCTGTAGGCTTACCTGCGCCCTCAACAACCTGCGGTACAGC 2221
Db 213 CTGTAGCTTGTGAGCAGGAGCTGTAGGCTTACCTGCGCCCTCAACAACCTGCGGTACAGC 154
QY 2222 GCGCTCCGCGAGGAGAACTTGAATCTGAGAACTTCCACACCTGAGGAGGAGGAGGAGG 2281
Db 153 GCGCTCCGCGAGGAGAACTTGAATCTGAGAACTTCCACACCTGAGGAGGAGGAGGAGGAGG 94
QY 2282 GAACCTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2341
Db 93 GAACCTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 34
QY 2342 CAGAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2374
Db 33 CAGAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1

RESULT 4
LOCUS BX460000 1201 bp mRNA linear EST 22-MAY-2003
DEFINITION BX460000 Homo sapiens FETAL BRAIN Homo sapiens CDNA clone

ACCESSION	CG6DF009YF05 5-PRIME, mRNA sequence.
VERSION	BX460000
KEYWORDS	BX460000.1 GI:31023191
SOURCE	EST.
ORGANISM	Homo sapiens (human)
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 1201)
AUTHORS	Li, W.-B., Gruber, C., Jesse, J. and Polayes, D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)
COMMENT	Contact: GenomeScope

FEATURES
cg1-ling-cluster:cg12peeg-CS0DF009C03QPL&ic1=661.f. Contact :
Pang Ling Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DF009C03QPL.
Location/Qualifiers

```

1. .1201
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="CSODP009YP05"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/note="Organ: brain; Vector: pCMVSORT 6; 1st strand cDNA
was primed with a NotI-oligo (dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSORT 6
vector. Library was not normalized."

```

ORIGIN

Query Match	39.1%;	Score 939.2;	DB 13;	length 1201;
Best Local Similarity	96.4%;	Pred. No. 1.2e-149;		
Matches 972; Conservative	8;	Mismatches 26;	Indels 2;	Gaps 2;

[illegible]

Db	463	CACGCGCTCGCGCCCTTTAGCTTACTCTACAGAGCGGA -	SeCGAGATTAACAGGACTCGACGC	521
Qy	515	GCCCTTCTAGCTGGCTAGGCTGGGCACTGGGGAGCCCA	CGAGCGAGCGAGCGAGCGAGCGG	574
Db	522	GCTTTCCTACGTGCGCTAGGCTGGGACTGGGGACCCCA	CGGCGGAGCGAGCGGCGAGCGG	581
Qy	575	AGCGCTGAGAGGCGAGCGGCGAGCGCCCGGAGCTCGA	AGATGCAACCGGCTGGAAACCGG	634
Db	582	AGCGCTGAGAGGCGAGCGGCGAGCGCCCGGAGCCGGA	AGATGCAACCGGCTGGAAACCGG	641
Qy	635	GCGAGTTTCCCGGAGGGGACGGGCGGCGAGAGTGGAG	AGGCGGCGCCCTCTGTCA	694
Db	642	GCGAGTTTCCCGGAGGGGACGGGCGGCGAGAGTGGAG	AGGCGGCGCCCTCTGTCA	701
Qy	695	CCTGAGCAACTGTGCGCTGCTCTCCCGCAGATTTCGTG	ACTGTGCTGTGATTC	754
Db	702	CCTGAGCAACTGTGCGCTGCTCTCCCGCAGATTTCGTG	ACTGTGCTGTGATTC	761
Qy	755	GAGGTGCGCAAGGCGGCGCTGCGCACTGTGTGCTTGTG	CCACCGGCTTGCGCGCGGAGCCCC	814
Db	762	GAGGTGCGCAAGGCGGCGCTGCGCACTGTGTGCTTGTG	CCACCGGCTTGCGCGCGGAGCCCC	821
Qy	815	CTGCTGCACATGCTCCCGCAGTGTGGGGGCGGCGCGG	CGTGTGCTGCGGCGCAGAGTTTCTG	874
Db	822	CTGCTGCACATGCTCCCGCAGTGTGGGGGCGGCGCGG	CGTGTGCTGCGGCGCAGAGTTTCTG	881
Qy	875	GAGTCCCTGAGCGCGGACTTGCCCGCCTTGAGAGCCAT	GGGACTTCACCTGTGGGCTGCA	934
Db	882	GAGTCCCTGAGCGCGGACTTGCCCGCCTTGAGAGCCAT	GGGACTTCACCTGTGGGCTGCA	941
Qy	935	GCGCCAGGAAACCAACCTGCTGGAATTAGCAATTGCTG	TGTAAGTTC - GCTGAGT	993
Db	942	GCGCCAGGAAACCAACCTGCTGGAATTAGCAATTGCTG	TGTAAGTTC - GCTGAGT	1001
Qy	994	GGATGGGCGAGTGCAGGATCACTCTCTTCCCGCAGAGCA	TAAACAGA 1041	
Db	1002	GGATGGGCGAGTGCAGGATCACTCTCTTCCCGCAGAGCA	TAAACAAA 1049	

[illegible]

```
FEATURES
SOURCE
    location/Qualifiers
    1..1080
    /organism="Homo sapiens"
    /mol_type="rRNA"
    /db_xref="taxon:9606"
```


Db	631	GGCCTGGGGAGTGTTCCTTCAACACTGGGGACCTGCTGTCGCAATACCAAGTTC	572
QY	1812	TCCGCTTCATGATGTCGAGACACCTTCAAGTGAAGGGGAGATGTGGCCACA	1871
Db	571	TCCGCTTCATGATGTCGAGACACCTTCAAGTGAAGGGGAGATGTGGCCACA	512
QY	1872	CCGAGGTGGCAGAGGCTTCGAGGCCCTAGATATTTCTTCAGAGGTGAACCTCATGAG	1931
Db	511	CCGAGGTGGCAGAGGCTTCGAGGCCCTAGATATTTCTTCAGAGGTGAACCTCATGAG	452
QY	1932	TCACGTGCCACGGGCAATGAAGGCGGCTGGAATGCAAGCCCTAGTTCGTGTC	1991
Db	451	TCACGTGCCACGGGCAATGAAGGCGGCTGGAATGCAAGCCCTAGTTCGTGTC	392
QY	1992	AACGCTTGGACCTTATGCAAGCTCTACACCAAGTCTGAGAACTTGCACCTATGACC	2051
Db	391	AACGCTTGGACCTTATGCAAGCTCTACACCAAGTCTGAGAACTTGCACCTATGACC	332
QY	2052	GGCCCCGATTCCTCAGGCTCCAGAGATCTTTGGCCACACAGAGACCTTCAACAGCAGA	2111
Db	331	GGCCCCGATTCCTCAGGCTCCAGAGATCTTTGGCCACACAGAGACCTTCAACAGCAGA	272
QY	2112	AAGTTCGATGGCAATGAGGGCTTCAGCCCGAGACCTGTGTAACCACTGTACGTT	2171
Db	211	AAGTTCGATGGCAATGAGGGCTTCAGCCCGAGACCTGTGTAACCACTGTACGTT	212
QY	2172	TGGACACAGCTGTAGGTCCTACCTGCCCTTCACACTGACCCTGGTACACGCGCTCTGG	2231
Db	211	TGGACACAGCTGTAGGTCCTACCTGCCCTTCACACTGACCCTGGTACACGCGCTCTGG	152
QY	2232	CAGGAAACCTTGGAATCTGAGAACTTCACACCTTAGGCACTGTAGAGAGAACTGTG	2291
Db	151	CAGGAAACCTTGGAATCTGAGAACTTCACACCTTAGGCACTGTAGAGAGAACTGTG	92
QY	2292	GGGTGGGGCCGCTTGCAGGTGTAAGTGGCTGTCAAGGATCTTTTCTATPACAGAACTGG	2351
Db	91	GGGTGGGGCCGCTTGCAGGTGTAAGTGGCTGTCAAGGATCTTTTCTATPACAGAACTGG	33
QY	2352	GTCACATATTTTGTATTAATGAGGCTGGAG	2381
Db	32	GTASATKTTT-TAATTAATGTGGCGGAG	4

RESULT 7	994 bp	mRNA	linear	EST 05-MAY-2003
LOCUS	EX354149			
DEFINITION	EX354149 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens			
ACCESSION	CDNA Clone CS0DC015YH21 5-PRIME, mRNA sequence.			
VERSION	EX354149			
KEYWORDS	EX354149.1 GI:30375843			
SOURCE	EST.			
ORGANISM	Homo sapiens (human)			
REFERENCE	Homo sapiens			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
JOURNAL	1. (bases 1 to 994)			
COMMENT	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.			
	Full-length cDNA libraries and normalization			
	Unpublished (2001)			
	Contact: Genoscope			
	Genoscope - Centre National de Sequencage			
	BP 191 91006 EVRY cedex - France			
	Email: secre@genoscope.cns.fr Web: www.genoscope.cns.fr			
	library was constructed by Life Technologies, a division of			
	Invitrogen. This sequence belongs to sequence cluster 661.f For			
	more information about this cluster, see			
	http://www.genoscope.cns.fr/			
	cgl.biml.ucsf.edu/cgl?seq=CS0DC015CD110P1&cluster=661.f . Contact :			
	Peng Liang Email: fliang@life.techn.com URL :			
	http://fulllength.invitrogen.com/ Invitrogen Corporation 1600			
	Faraday Avenue Genoscope sequence ID : CS0DC015CD110P1.			
FEATURES	Location/Qualifiers			

```

source
1..394
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="CSDDC015HR21"
/issue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMSPORT 6 vector. Library was normalized."

```

[illegible]

High quality sequence stop: 686.
Location/Qualifiers

FEATURES
source

```

1. 908
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6091179"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DR10B (phage-resistant)"
/clone_lib="NIH_MGC_112"
/notes="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: This is a
NIH_MGC Library."

```

ORIGIN

Query Match	34.9%	Score 839.6;	DB 13;	Length 908;
Best Local Similarity	98.6%;	Pred. No. 9.4e-133;		
Matches 868;	Conservative 0;	Mismatches 9;	Indels 3;	Gaps 2;

QY	1502	TCACGTACACAGACAGCGGAGCGCTGTGGGAGCGTCTTCTCTGGCTTTCACAGCATATCT	1582
Db	1	TCACGTACACAGACAGCGGAGCGCTGTGGGAGCGTCTTCTCTGGCTTTCACAGCATATCT	60
QY	1563	TCCTCTTCTCTGATTCGTATGTACACACAGAGAGCCAAATTCCGACATCCCCACAG	1622
Db	61	TCCTCTTCTCTGATTCGTATGTACACACAGAGAGCCAAATTCCGACATCCCCACAG	120
QY	1623	GGACGTATAGGACCATCTCCAGGTGAGCCAGGGCTGTGCTGTGGCCCGGTAAAGCACAG	1682
Db	121	GGACGTATAGGACCATCTCCAGGTGAGCCAGGGCTGTGCTGTGGCCCGGTAAAGCACAG	180
QY	1683	AGTCCCATTTCTGSGCTATGCTGCGGGCCAGAGCTTGCCACAGGSGAAAGTTGCTPAAAG	1742
Db	181	AGTCCCATTTCTGSGCTATGCTGCGGGCCAGAGCTTGCCACAGGSGAAAGTTGCTPAAAG	240
QY	1743	ATGCTTCCGGCTGSGGATGTTTTCTTCAACAATCTGGGAGCCGTGCTGCTCGAGTACAG	1802
Db	241	ATGCTTCCGGCTGSGGATGTTTTCTTCAACAATCTGGGAGCCGTGCTGCTCGAGTACAG	300
QY	1803	AAGTTTTCTCGCTTCATGATGTACTGAGACACCTTCAGGTGAAAGSGGAGAAATG	1862
Db	301	AAGTTTTCTCGCTTCATGATGTACTGAGAGACCTTCAGGTGAAAGSGGAGAAATG	360
QY	1863	TGGCCACACAGAGTGGACAGAGCTTCCAGCGCCCTAGATTTCTTCAAGAGAGTGAAG	1922
Db	361	TGGCCACACAGAGTGGAGAGGCTTCCAGCGCCCTAGATTTCTTCAAGAGAGTGAAG	420
QY	1923	TCATAGATCTACTGTGCAGGACATGAAGCAGAGGCTGGATATGAGACCTTAAGTTCTGC	1982
Db	421	TCATAGATCTACTGTGCAGGACATGAAGCAGAGGCTGGATATGAGACCTTAAGTTCTGC	480
QY	1983	GTCCCCCCACGCTTTGGACCTTATGACAGCTTACACCAACGTGTCTGAGACTTGGCCAC	2042
Db	481	GTCCCCCCACGCTTTGGACCTTATGACAGCTTACACCAACGTGTCTGAGACTTGGCCAC	540
QY	2043	CTATATCCCGGCCCCGATTTCTCAGGCTCCAGAGCTTTTGAGCCACCAACAGACCTTCA	2102
Db	541	CTATATCCCGGCCCCGATTTCTCAGGCTCCAGAGCTTTTGAGCCACCAACAGACCTTCA	600
QY	2103	AACAGCAAAAGTTGGATGGCAAAATGAGGGCTTGACCCACAGACCCCTGTCTGACCCAC	2162
Db	601	AACAGCAAAAGTTGGATGGCAAAATGAGGGCTTGACCCACAGACCCCTGTCTGACCCAC	660
QY	2163	TGTACCTTTGAGACCAAGCTGTAGGTGTCTTACCTGACCCCTCACAACCTGCCCCGTACAGG	2222
Db	661	TGTACCTTTGAGACCAAGCTGTAGGTGTCTTACCTGACCCCTCACAACCTGCCCCGTACAGG	720
QY	2223	CCCTCTGACAGAAACCTTGCATCTGAGAACTTCCACACTTGAGGACCTGAGAGAG	2282

Db 721 CCTCTCTGGCAGAAACCTTGGAACTTGGAGAACTTCCACACTGAGGACACTGAGAGAGG 780

Oy 2283 AACTCTGTGGGGGT-GGGGCGCTGTGCAGTGTACTGGGCTGTCAAGGATCTTTCTATAC 2344

Db 781 AACTCTGTGGGGGTGGGGGCGCGTTGCAGGTGTACTGGGCTGTCAAGGATCTTTCTATAC 840

```

QY      2342 CAGAACTGCGG--TCACTATTTTGTATTAATGTGGCTGG 2375
          |||||
DB      841  CCGAACTGCGGGTCACTATTTTGTATTAATGTGGGTGG 880

```

RESULT 10	870 bp	MRNA	linear	EST 23-MAY-2003
AL589216/c				
LOCUS				
DEFINITION				
AL589216 Homo sapiens				
NEUROBLASTOMA CRT 50-NORMALIZED				
Homo sapiens				
CDNA clone CS0DD0021503 3'-PRIME, mRNA sequence.				

FEATURES

၇၉

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DD0021E03"
/tissue_type="NEUROBLASTOMA COT 50-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (drr)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

```

ORIGIN

Query Match	34.8%	Score 837	DB 9	Length 870
Best Local Similarity	98.2%	Pred. No. 2.6e-132		
Matches 855	Conservative 11	Mismatches 3	Indels 2	Gaps 2

OY	1485	AGGGCACTGTGCCACCACTCAACTACACA	-GAGACAGGGGGGCGCTGTGGGGGTGCTTC	1545
Db	870	AGGGCAACTGTGCCACCACTCAACTACACA	AGGGAGAGGGGGCGCTGTGGGGGTGCTTC	811
OY	1544	TGGCTTACAAACCATATTTCCCTTCCTTGATTTGGCTATATATGTCACACAGAGAG		1603
Db	810	TGGCTTDAACAAACCATATTTCCCTTCCTTGATTTGGCTATATATGTCACACAGAGAG		751
OY	1604	CCAATTGGGAGACCCCAAGGCACTGTATGGCCACATCTCCAGGTAGAGCAGAGGCTGCTG		1663
Db	750	CCAATTGGGAGACCCCAAGGCACTGTATGGCCACATCTCCAGGTAGAGCAGAGGCTGCTG		691
OY	1664	GTGGCCCCGGTAGCCACAGTCCCATTTCTTGGGCTATCTGGCGGGCCAGAGCTGACC		1722

Db 690 GTGGCCCGGTAAGCAGACATCCCATTCCTGGGCTATGCTGGC-GGCCAGACTGGCC 632
 QY 1724 CAGGGGAAGTGTCTAAAGAGATGTCTTCGGGCTGGGAGTGTCTTTCACACTGAGGAC 1783
 Db 631 CAGGGGAAGTGTCTAAAGAGATGTCTTCGGGCTGGGAGTGTCTTTCACACTGAGGAC 572
 QY 1784 CTGGGCTGTCCGATGACCAAGTCTTTCGGGCTTCCATGATGCTACTGAGACACTTC 1843
 Db 571 CTGGGCTGTCCGATGACCAAGTCTTTCGGGCTTCCATGATGCTACTGAGACACTTC 512
 QY 1844 AGTGAAGGGGAGATGTG3CCACAACCGAGGTGACAGAGTCTTCAGAGCCCTAGAT 1903
 Db 511 AGTGAAGGGGAGATGTG3CCACAACCGAGGTGACAGAGTCTTCAGAGCCCTAGAT 452
 QY 1904 TTTCTTCAGAGAGTGAAGCTCTATGAGTCACTGTGCCAGGGCATGAAAGGAGGCTGGA 1963
 Db 451 TTTCTTCAGAGAGTGAAGCTCTATGAGTCACTGTGCCAGGGCATGAAAGGAGGCTGGA 392
 QY 1964 ATGGCAGCCCTAGTCTGCTGCCGCCGCCCGCTTGGACCTTATGACGCTTACACCCAC 2023
 Db 391 ATGGCAGCCCTAGTCTGCTGCCGCCGCCCGCTTGGACCTTATGACGCTTACACCCAC 332
 QY 2024 GTGTCTGAAGACTTCCACTTATATGCCGCCCGGCTTCTCAGGCTTCCAGAGTCTTTG 2083
 Db 331 GTGTCTGAAGACTTCCACTTATATGCCGCCCGGCTTCTCAGGCTTCCAGAGTCTTTG 272
 QY 2084 GGCACACAGAGACCTTAAACAGAGAAAGTGTGGATGAGCAATGAGGGCTTCCAGACCC 2143
 Db 271 GGCACACAGAGACCTTAAACAGAGAAAGTGTGGATGAGCAATGAGGGCTTCCAGACCC 212
 QY 2144 AGCACCCTGTCTGAACCACTGTAGCTTCTGACACAGGCTGAGTGTGCTACTGACCCCTC 2203
 Db 211 AGCACCCTGTCTGAACCACTGTAGCTTCTGACACAGGCTGAGTGTGCTACTGACCCCTC 152
 QY 2204 ACAACTGCCCGGTACAGGGCCCTCTCTGGCAGAGAACTTGAATCTGAAGCTTCCACAC 2263
 Db 151 ACAACTGCCCGGTACAGGGCCCTCTCTGGCAGAGAACTTGAATCTGAAGCTTCCACAC 92
 QY 2264 CTGAGGCACTGAGAGAGAACTCTGTGGGCTGGGGCCGTGAGGCTGATCTGGGCTGT 2223
 Db 91 CTGAGGCACTGAGAGAGAACTCTGTGGGCTGGGGCCGTGAGGCTGATCTGGGCTGT 32
 QY 2324 CAGGATCTTTTCTATACCAAGACTGCGGTC 2354
 Db 31 CAGGATCTTTTCTATACCAAGATMGCTM 1

RESULT 11
 AL521330
 LOCUS 934 bp mRNA linear EST 22-MAY-2003
 DEFINITION AL521330 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens
 CDNA clone CS0DB001YE18 5-PRIME, mRNA sequence.
 ACCESSION AL521330
 VERSION AL521330.2 GI:31039633
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 934)
 L1.W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On Feb 13, 2001 this sequence version replaced gi:12784823.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 661.f. For more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DB001BC09Q0Picluster=661.f. Contact :

Feng Liang Email: fliang@life.techn.com URL: <http://fulllength.invitrogen.com/InvitrogenCorporation1600FaradayAvenueGenoscopeSequenceID:CS0DB001BC09Q0P1>
 Location/Qualifiers
 1. 934
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DB001YE18"
 /issue="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"
 /clone_lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"
 /note="Full length cDNA was primed with a NotI-clisodIT primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
 Query Match 34.7%; Score 833.6; DB 9; Length 934;
 Best Local Similarity 98.6%; Pred. No. 16-111; 7; Indels 2; Gaps 2;
 Matches 858; Conservative 3; Mismatches 77;

QY 11 CGGGTTCTGCTCTCCGCCCGGTGTGAGTGTGTGGGGGCTG3GTGGAAATG3CGGTGTC 70
 Db 67 CGGGTTCTGCTCTCCGCCCGGTGTGAGTGTGTGGGGGCTG3GTGGAAATG3CGGTGTC 126
 QY 71 CAGGCAAGCGGCGTCCCTGGAAGAGAGTCTCAGCTGAGACGAGCGGCGCTAGGTTT 130
 Db 127 CAGGCAAGCGGCGTCCCTGGAAGAGAGTCTCAGCTGAGACGAGCGGCGCTAGGTTT 186
 QY 131 CCGAAGGAGAGATCAAGGATGTTTGGAGCGGCTGGAACAGACGCTGCGATAGAGAA 190
 Db 187 CCGAAGGAGAGATCAAGGATGTTTGGAGCGGCTGGAACAGACGCTGCGATAGAGAA 246
 QY 191 GCGGCTCCATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 250
 Db 247 GCGGCTCCATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 306
 QY 251 CTGAAGCTACACTCTGAGCGCAGATTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 310
 Db 307 CTGAAGCTACACTCTGAGCGCAGATTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 366
 QY 311 CCGAGCTGTGCTGCAAAAGGCTCTTGAAGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTG 370
 Db 367 CCGAGCTGTGCTGCAAAAGGCTCTTGAAGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTG 426
 QY 371 CCGAAGAGTCCGAGAGGGGGCTGAGCGCTGAGCGCTGAGCGCTGAGCGCTGAGCGCTGAG 430
 Db 427 CCGAAGAGTCCGAGAGGGGGCTGAGCGCTGAGCGCTGAGCGCTGAGCGCTGAGCGCTGAG 486
 QY 431 CGGCGCGCAGACCTTTCTCATTTCAAGCTGCGCGCTGTTAGTACTCAGAGCGGAG 490
 Db 487 CGGCGCGCAGACCTTTCTCATTTCAAGCTGCGCGCTGTTAGTACTCAGAGCGGAG 546
 QY 491 CGGAGAGTAAAGAGGCTGCAAGCGCTCTTCTCAAGTGTGCTGAGCTGAGCTGAGCTGAG 550
 Db 547 CGGAGAGTAAAGAGGCTGCAAGCGCTCTTCTCAAGTGTGCTGAGCTGAGCTGAGCTGAG 606
 QY 551 GACGCGCGAGCAGAGCGGAGCGCTGAGAGAGCGAGCGGAGCGGAGCGGAGCGGAG 610
 Db 607 GACGCGCGAGCAGAGCGGAGCGCTGAGAGAGCGAGCGGAGCGGAGCGGAGCGGAG 666
 QY 611 GGAAGTACAGCGGCGGAGAGCGGCGGAGCTTGGCGGAGGAGAGCTGCGCGAGAGT 670
 Db 667 GGAAGTACAGCGGCGGAGAGCGGCGGAGCTTGGCGGAGGAGAGCTGCGCGAGAGT 726
 QY 671 GAGAGAGCGGCGGCGCTGCTGCACTGAGAGCACTGTGAGCTGCTGCTGCTGCTGCTG 730
 Db 727 GAGAGAGCGGCGGCGCTGCTGCACTGAGAGCACTGTGAGCTGCTGCTGCTGCTGCTG 786
 QY 731 CCGAGATTTCTGTGCTGTGTTGCGGCTGCGCAAGCGCGGCTGCGCACTGCTGCTG 790
 Db 787 CCGAGATTTCTGTGCTGTGTTGCGGCTGCGCAAGCGCGGCTGCGCACTGCTGCTG 846

Qy	791	CCACACGCGCCTTGAGCGCGGAGCCCGCGTGTGACCTGCTCCGAGCTGCGGCGCGCGCGC	850
Db	847	CCACACGCGCCTTGAGCGCGGAGCCCGCGTGTGACCTGCTCCGAGCTGCGGCGCGCGCGC	905
Qy	851	CTGCTGCTGCGCCCAAGATTCTTGAGTGC	880
Db	906	TGGGTGCTGGCG-CARASTTCTTGAGTGC	934

RESULT	12			
LOCUS	B0678639			
DEFINITION	BOE78639	903 bp	mRNA	linear
	AGNCNCOURT_8192626	NIH_MGC_112	Homo sapiens	cDNA clone IMAGE:626111
	5', mRNA sequence.			
ACCESSION	BOE78639			
VERSION	BOE78639.1	GI:21791330		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 903)			
AUTHORS	NIH-MGC	http://mgs.nci.nih.gov/.		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)			
JOURNAL	unpublished (1998)			
COMMENT	Contact: Robert Strausberg, Ph.D.			

FEATURES

```

1..903
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:66281109"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NH1 MGC_112"
/notes="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using Zap-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NH1 MGC Library."

```

ORIGIN

	Query Match	Similarity	Score	DB	Length
Best Local Similarity	99.1%	32.7%	766.8	DB 13	903
Matches	791	Conservative	0	Mismatches	7
				Indels	0
				Gaps	0

	Query	1503	TCAACTACAGAGACGCGGGCGCTGTGCGGCGCTTCTGCGCTTTACAGCAATCT	1562
Db	1	TCAACTACAGAGACGCGGGCGCTGTGCGGCGCTTCTGCGCTTTACAGCAATCT	60	
Qy	1563	TCCCCTTCCTTGATTCGCTATGATGTCACCAAGAGGCCAATTGGGACCCCCAGG	1622	
Db	61	TCCCCTTCCTTGATTCGCTATGATGTCACCAAGAGGCCAATTGGGACCCCCAGG	120	
Qy	1623	GGAATGTATGGCCACATCTCCAGGTGAAGCCAGGGCTGTGTGGCCCGGTAACCAAGC	1682	
Db	121	GGAATGTATGGCCACATCTCCAGGTGAAGCCAGGGCTGTGTGGCCCGGTAACCAAGC	180	
Qy	1683	AGTCCCATTCCTGGGCTATGCTGCGGGCCAGAGCTGGCCAGGAGGAAATTGCTAAAG	1744	

Db	181	AGTCCCATTTCTGCGGCTATATGCTCGGCGGCGCAGAGCTGGGCCACAGGGGAAGTTGCTAAAGG	240
Qy	1743	ATGTCCTTCGGGCTCTGGGGATGTATTTCTTTCAACAATCTGGGGACCTGCTGTCTGGCATGACC	1802
Db	241	ATGCTCTTCGGGCTCGGGGATGTTTCTTCAACACTGGGAGACTGTCTGGTCTGCATGAC	300
Qy	1803	AAGGTTTTCTCGGCTTCATGATGTAATGTAAGACACCTTCAGGTGTGAGAGGGGAGATG	1862
Db	301	AAGGTTTTCTCGGCTTCATGATGTAATGTAAGACACCTTCAGGTGTGAGAGGGGAGATG	360
Qy	1863	TGGCCACAACCGAGGTGTGCAGAGGTCTTTCGAGGCCCTTAGATTTTTCTTCAGAGGTGAACG	1922
Db	361	TGGCCACAACCGAGGTGTGCAGAGGTCTTTCGAGGCCCTTAGATTTTTCTTCAGAGGTGAACG	420
Qy	1923	TCTATGAGATCACTGTGTGCAGGCGATGAAGCAGGGCTGGAATGGACCCCTAGTTCTGC	1982
Db	421	TCTATGAGATCACTGTGTGCAGGCGATGAAGCAGGGCTGGAATGGACCCCTAGTTCTGC	480
Qy	1983	GTCCCGCCCAACGCTTGTGAGCCTTATGACAGCTCTACACCCACGTTCTGAGAACTTGCCAC	2042
Db	481	GTCCCGCCCAACGCTTGTGAGCCTTATGACAGCTCTACACCCACGTTCTGAGAACTTGCCAC	540
Qy	2043	CTTATGCCCCGCCCCGATTTCTCAAGGCTCCAGAGTCTTTGGCCACACACAGAACCTTCA	2102
Db	541	TTTATGCCCCGCCCCGATTTCTCAAGGCTCCAGAGTCTTTGGCCACACACAGAACCTTCA	600
Qy	2103	AACAGCAAAAGTTCCGATGGCAATGAGGACTTCGACCAGCACCCCTGTCTAACCCAC	2162
Db	601	AACAGCAAAAGTTCCGATGGCAATGAGGACTTCGACCAGCACCCCTGTCTAACCCAC	660
Qy	2163	TGTACGTTCTTGAGACCAAGCTGTAGGTGGCTTACTGCCCCCTACCACTGCCCCGGTACAGCG	2222
Db	661	TGTACGTTCTTGAGACCAAGCTGTAGGTGGCTTACTGCCCCCTACCACTGCCCCGGTACAGCG	720
Qy	2223	CCCTCTCTGGCAGGAAACTTTGGAATCTGAGAACTTCCACACTCTGAGGACCTGTGAGAGAGG	2282
Db	721	CCCTCTCTGGCAGGAAACTTTGGAATCTGAGAACTTCCACACTCTGAGGACCTGTGAGAGAGG	780
Qy	2283	AACTCTGTGGGGTGTGGGG 23 00	
Db	781	AACTCTGTGGGGTGTGGGG 798	

RESULT 13
AL529217 873 bp mRNA linear EST 23-MAY-2003
LOCUS AL529217
DEFINITION Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
CDNA clone CS0DD002YE03 5-PRIME, mRNA sequence.
ACCESSION AL529217
VERSION AL529217.2 GI:31067061
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 873)
AUTHORS Li W.B., Gruber C., Jesssee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
COMMENT Unpublished (2001).
On Feb 13, 2001 this sequence version replaced gi:12792710.

FEATURES

source 1..873

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CSDD002YB03"

/issue_type="NEUROBLASTOMA COT 50-NORMALIZED"

/clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-clig(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 32.2%; Score 774.8; DB 9; Length 873;

Best Local Similarity 98.8%; Pred. No. 1e-121;

Matches 809; Conservative 2; Mismatches 5; Indels 3; Gaps 3;

101 TCTCAGCTAGAAAGAGCGGCTTCTAGGTTTTCGAAAGAGATCAGGAGTTTTCGAGC 160

49 TCTCAGCTAGAAAGAGCGGCTTCTAGGTTTTCGAAAGAGATCAGGAGTTTTCGAGC 108

161 GGCCTGAAACGAGGCTGCGATAGAGAGAGGCTCCATGCTGCTGCTGCTGCTG 220

109 GGCCTGAAACGAGGCTGCGATAGAGAGAGGCTCCATGCTGCTGCTGCTGCTG 168

221 CCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 280

169 CCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 228

281 TGGCTTCGCGGAGATTTGGCTTTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 340

229 TGGCTTCGCGGAGATTTGGCTTTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 288

341 GCTCGCGGCTTGGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 400

289 GCTCGCGGCTTGGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 348

401 GCTCGCGGCTTGGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 460

349 GCTCGCGGCTTGGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 407

461 TCGCGGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 520

408 TCGCGGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 466

521 CTAAGTGGCTAGAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 580

467 CTAAGTGGCTAGAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 526

581 GGAAGAGGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 640

527 GGAAGAGGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 586

641 TTTGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 700

587 TTTGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 645

701 GCAACTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 760

646 GCAACTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 705

761 GCAAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 820

706 GCAAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 765

821 CACTGCTCTCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 880

766 CACTGCTCTCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 825

881 CTGAGAGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 940

826 CTGAGAGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 864

RESULT 14

BI820310 1044 bp mRNA linear EST 04-OCT-2001

LOCUS 603036876F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5177899 5',

DEFINITION mRNA sequence.

ACCESSION BI820310.1 GI:15931860

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

NIH-MGC http://mgc.nci.nih.gov/.

1 (bases 1 to 1044)

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgaabs-remail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)

DNA Sequencing by: Inocyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at: http://image.llnl.gov

Plate: LIML1443 row: g column: 20

High quality sequence stop: 789.

Location/Qualifiers

1..1044

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5177899"

/lab_host="DH10B"

/clone_lib="NIH_MGC_115"

/note="Organ: Pooled brain, lung, testis; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 31.9%; Score 767.2; DB 12; Length 1044;

Best Local Similarity 92.6%; Pred. No. 2.1e-120;

Matches 906; Conservative 0; Mismatches 58; Indels 14; Gaps 9;

122 CTAGCTTTTCGAGAGGAGATCAGGAGTTTTCGAGAGGAGTGGAGACGAGCGGCGG 181

1 CTAGGTTTTCGAGAGGAGATCAGGAGTTTTCGAGAGGAGTGGAGACGAGCGGCGG 60

182 ATTAGAGAAAGCGGCTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 241

61 ATTAGAGAAAGCGGCTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120

242 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 301

121 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180

302 TTTGCGGTGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 361

181 TTTGCGGTGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240

362 GCGCGCGAGCGGAGAGTCCGAGAGGAGGCTGCACTGCTGCTGCTGCTGCTGCTGCTG 421

241 GCGCGCGAGCGGAGAGTCCGAGAGGAGGCTGCACTGCTGCTGCTGCTGCTGCTGCTG 300

422 GCCCAGACGCGCGCGCAGACCTTTCATTCAAGGCTGCGGCGCTTAGCTACTCA 481

Wed Jul 7 12:07:09 2004

us-10-030-226-1.rst

Page 15

Db 847 AG 848

Search completed: July 3, 2004, 03:46:26
Job time : 617 secs

This Page Blank (uspto)

CC gene encoding it are useful as targets for the treatment or prevention of
CC diseases associated with defective metabolism of long-chain fatty acids.
CC Such diseases include cardiomyopathy, skeletal muscle disorders, or renal
CC failure. PBC67 nucleic acids may also be used in gene therapy for such
CC disorders. The present sequence represents cDNA encoding PBC67
XX

Sequence 2405 BP; 410 A; 735 C; 781 G; 479 T; 0 U; 0 Other;

Query Match 100.0%; Score 2405; DB 4; Length 2405;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GCACTCTCCGGGTTTCTGCTCTCCGCCCTGTGTGAGTGTGGGGGCTTGGTGGGAT 60
DB 1 GCACTCTCCGGGTTTCTGCTCTCCGCCCTGTGTGAGTGTGGGGGCTTGGTGGGAT 60
QY 61 GGGCGTGTCCAGACGCGCGCTCCCTGGAAGAGAGTCTCAGCTAGAACGACGCGC 120
DB 61 GGGCGTGTCCAGACGCGCGCTCCCTGGAAGAGAGTCTCAGCTAGAACGACGCGC 120
QY 121 CCTAGGTTTCCGAAAGGAGATCAAGGATGTTTGCAGCGGCTGMAACCAAGCGTCC 180
DB 121 CCTAGGTTTCCGAAAGGAGATCAAGGATGTTTGCAGCGGCTGMAACCAAGCGTCC 180
QY 181 GATPAGAGAGAGGGGCTCCATGAGTGCCTCTGCTGCTGACCTCTGCTGCTGCTTACC 240
DB 181 GATPAGAGAGAGGGGCTCCATGAGTGCCTCTGCTGCTGACCTCTGCTGCTGCTTACC 240
QY 241 GCTGCTGCTGCTGAACTACCTCTGCGCGAGTTGCGCTTGGCTTCCGCGGACTTGGC 300
DB 241 GCTGCTGCTGCTGAACTACCTCTGCGCGAGTTGCGCTTGGCTTCCGCGGACTTGGC 300
QY 301 CTTTGGGTTGCGAGCTCTGTGCTGCAAAAGGGCTCTTGAGTCTCGGCCCTGCGCGAGC 360
DB 301 CTTTGGGTTGCGAGCTCTGTGCTGCAAAAGGGCTCTTGAGTCTCGGCCCTGCGCGAGC 360
QY 361 TGGCGCGACCCGGAAGTCCCGAGGGGGCTGCAAGCTGAGCTGAGCTTGGCGGAACT 420
DB 361 TGGCGCGACCCGGAAGTCCCGAGGGGGCTGCAAGCTGAGCTGAGCTTGGCGGAACT 420
QY 421 GGGCCAGACGGGGCGGCGGCACTTTCTCATTCAGGCTGTGGGGCTTTAGTATTC 480
DB 421 GGGCCAGACGGGGCGGCGGCACTTTCTCATTCAGGCTGTGGGGCTTTAGTATTC 480
QY 481 AAGAGGCGAGCGCGAGAGTACAGGGCTCAAGCGCTTCTACCTGAGCTAGCTGGA 540
DB 481 AAGAGGCGAGCGCGAGAGTACAGGGCTCAAGCGCTTCTACCTGAGCTAGCTGGA 540
QY 541 CTGGGAGACCCGACGGCGGCGACAGCGGCGAGGAGCGCTGAGAGAGCGCGGACGC 600
DB 541 CTGGGAGACCCGACGGCGGCGACAGCGGCGAGGAGCGCTGAGAGAGCGCGGACGC 600
QY 601 GCGGAGAGCGGAGATGACAGCGGCGAGAAAGCGGAGATTGCGGAGGAGAGCGTGC 660
DB 601 GCGGAGAGCGGAGATGACAGCGGCGAGAAAGCGGAGATTGCGGAGGAGAGCGTGC 660
QY 661 CGGACAGGTGAGAGAGCGCGCTCTGTCTGCTGACCTGAGCACTGTGGCGTCTCT 720
DB 661 CGGACAGGTGAGAGAGCGCGCTCTGTCTGCTGACCTGAGCACTGTGGCGTCTCT 720
QY 721 CCGCGGTGCGGAGATTCTGTGCTGTGGTTCGGGCTGCGCAAGCGCGCTCGGAC 780
DB 721 CCGCGGTGCGGAGATTCTGTGCTGTGGTTCGGGCTGCGCAAGCGCGCTCGGAC 780
QY 781 TGGCTTTGTGCCACCGGCTGTGCGCGGGGCGGCTGCTGCACTGCTCCGAGCTGCGG 840
DB 781 TGGCTTTGTGCCACCGGCTGTGCGCGGGGCGGCTGCTGCACTGCTCCGAGCTGCGG 840
QY 841 CGGCGCGGCTGTGCTGCGCGCAAGATTCTGAGATCTCTGAGACCGGACCTTCCGC 900
DB 841 CGGCGCGGCTGTGCTGCGCGCAAGATTCTGAGATCTCTGAGACCGGACCTTCCGC 900
QY 901 CCGAGAGCATGAGGCTCCACTGTGGGCTGAGGCGCAGGAACTCCTGTGGAAT 960
```

```
DB 901 CCGAGAGCATGAGGCTCCACTGTGGGCTGAGGCGCAGGAACTCCTGTGGAAT 960
QY 961 TAGGATTTGCTGGTGAAGTGTCCGCTGAAGTGAAGGCGGACAGTACCTCTC 1020
DB 961 TAGGATTTGCTGGTGAAGTGTCCGCTGAAGTGAAGGCGGACAGTACCTCTC 1020
QY 1021 TTCCCCCAGAGATTAACAGACGCTGCTGTACATCTTCACTCTGAGCAACGGGCT 1080
DB 1021 TTCCCCCAGAGATTAACAGACGCTGCTGTACATCTTCACTCTGAGCAACGGGCT 1080
QY 1081 CCCCAAGGCTGCTCGATCATCTGAAGATCTGCAATGCTCAAGGCTTCTATACGT 1140
DB 1081 CCCCAAGGCTGCTCGATCATCTGAAGATCTGCAATGCTCAAGGCTTCTATACGT 1140
QY 1141 GTGTGTGTCCACAGAGAGATGTATCTACTCGCCCTCCCACTTACCAATGTCCG 1200
DB 1141 GTGTGTGTCCACAGAGAGATGTATCTACTCGCCCTCCCACTTACCAATGTCCG 1200
QY 1201 TTCCCTGTGGGAGATGTGGGCTGCAATGGGACATGGGCGCACAGTGTGCTGAATCCA 1260
DB 1201 TTCCCTGTGGGAGATGTGGGCTGCAATGGGACATGGGCGCACAGTGTGCTGAATCCA 1260
QY 1261 GTTCTGGCTGTCAAGTTCTGGGAGATTGCCAGACACAGGCTGACGCTTCCAGTA 1320
DB 1261 GTTCTGGCTGTCAAGTTCTGGGAGATTGCCAGACACAGGCTGACGCTTCCAGTA 1320
QY 1321 CATTGGGAGCTGTGCGGATCTGTCAACAGCGCCGAGAGAGGACGAGAGCTGGCCA 1380
DB 1321 CATTGGGAGCTGTGCGGATCTGTCAACAGCGCCGAGAGAGGACGAGAGCTGGCCA 1380
QY 1381 TAAAGTCCGCTGCGAGTGGGACAGCGGGCTGGCCGACATCTGAGACGTTTGTGCG 1440
DB 1381 TAAAGTCCGCTGCGAGTGGGACAGCGGGCTGGCCGACATCTGAGACGTTTGTGCG 1440
QY 1441 GCGCTTGGGCGCCCTGCAAGTGTGTGAGACATATGACACAGAGGCAAGTGGCCAC 1500
DB 1441 GCGCTTGGGCGCCCTGCAAGTGTGTGAGACATATGACACAGAGGCAAGTGGCCAC 1500
QY 1501 CATCACTACACAGACAGCGGGCGCTGTGGGCTGTCTTCCGCTTTACAGACAT 1560
DB 1501 CATCACTACACAGACAGCGGGCGCTGTGGGCTGTCTTCCGCTTTACAGACAT 1560
QY 1561 CTTCCCTCTCTCTTGAATGCTATGATGTCAACACAGAGACCAATTCGGGACCCCA 1620
DB 1561 CTTCCCTCTCTCTTGAATGCTATGATGTCAACACAGAGACCAATTCGGGACCCCA 1620
QY 1621 GGGGCACTGTATGAGCAATCTCCAGGTGAGCGAGGCTGTGTGGCCCGGTAAGCCA 1680
DB 1621 GGGGCACTGTATGAGCAATCTCCAGGTGAGCGAGGCTGTGTGGCCCGGTAAGCCA 1680
QY 1681 GAGTCCCAATTCCTGAGCTATGCTGGCGGACAGCTGGCCACGGGAAATTGCTAA 1740
DB 1681 GAGTCCCAATTCCTGAGCTATGCTGGCGGACAGCTGGCCACGGGAAATTGCTAA 1740
QY 1741 GAGTCCCAATTCCTGAGCTATGCTGGCGGACAGCTGGCCACGGGAAATTGCTAA 1800
DB 1741 GAGTCCCAATTCCTGAGCTATGCTGGCGGACAGCTGGCCACGGGAAATTGCTAA 1800
QY 1801 CCAAGGTTTCTCCGCTTCATGATGTACTGAGACACTTCAAGTGAAGGGGAGAA 1860
DB 1801 CCAAGGTTTCTCCGCTTCATGATGTACTGAGACACTTCAAGTGAAGGGGAGAA 1860
QY 1861 TGTGGCCAAACGAGGTGCAAGGCTTTCGAGGCTTCAAGTTCCTCAGGAGTGA 1920
DB 1861 TGTGGCCAAACGAGGTGCAAGGCTTTCGAGGCTTCAAGTTCCTCAGGAGTGA 1920
QY 1921 CGTCTATGAGTCACTGTGCAAGGAGTGAAGGCGGCTGGAATGCAAGCCTTACTT 1980
DB 1921 CGTCTATGAGTCACTGTGCAAGGAGTGAAGGCGGCTGGAATGCAAGCCTTACTT 1980
QY 1981 GCGTCCCGCAAGCTTTGGAATTTATGACGCTCTACCCCAAGTGTCTGAAGACTTGGC 2040
```

DB 1981 GCGTCCCCCAGCTTTGACCTTATGACGCTTACACCAACGCTGTGAGAACTTGCC 2040
QY 2041 ACCTTATGCCCCCGCCGATTCCTAGAGCTCCAGAGCTTTTGGCCACACAGAGACTT 2100
DB 2041 ACCTTATGCCCCCGCCGATTCCTAGAGCTCCAGAGCTTTTGGCCACACAGAGACTT 2100
QY 2101 CAACAGCAGAAAGTTCGATGAGCAATGAGAGGCTTCGACCCAGCACTGTCTGACCC 2160
DB 2101 CAACAGCAGAAAGTTCGATGAGCAATGAGAGGCTTCGACCCAGCACTGTCTGACCC 2160
QY 2161 ACTGACGTTTGTGACCAAGCTGTAGTGTCTTACCTGCCCCCTCACACTGCCCCGATACG 2220
DB 2161 ACTGACGTTTGTGACCAAGCTGTAGTGTCTTACCTGCCCCCTCACACTGCCCCGATACG 2220
QY 2221 CGCCCTCTGGAGAGAAACCTTCGAACTGGAATCTTCACACCTGAGGACCTGAGAGA 2280
DB 2221 CGCCCTCTGGAGAGAAACCTTCGAACTGGAATCTTCACACCTGAGGACCTGAGAGA 2280
QY 2281 GGAACCTGTGTGGAGTGGGGGCGGTGCAAGGTGTACTGTGAGGATCTTTCTATA 2340
DB 2281 GGAACCTGTGTGGAGTGGGGGCGGTGCAAGGTGTACTGTGAGGATCTTTCTATA 2340
QY 2341 CGAAGACTGCGGTCTATTTTGTATATAAATGTGCTGTGAGACTGATCCAGCTGTCTGGA 2400
DB 2341 CGAAGACTGCGGTCTATTTTGTATATAAATGTGCTGTGAGACTGATCCAGCTGTCTGGA 2400
QY 2401 CCTAC 2405
DB 2401 CCTAC 2405

RESULT 2

ADA84114 ID ADA84114 standard; DNA; 2430 BP.

ADA84114;

20-NOV-2003 (first entry)

Human POM124 gene.

human; marker: expressed sequence tag; EST; arabidopsis; tumour;

stress-induced phenotype; hyperosmotic stress; colon cancer; immunogen;

vaccine; ds; gene.

Homo sapiens.

MO2002103028-A2.

27-DEC-2002.

30-MAY-2002; 2002MO-IB004189.

30-MAY-2001; 2001US-0293999P.

22-OCT-2001; 2001US-030457P.

19-FEB-2002; 2002US-0357144P.

(BIOM-) BIOMEDICAL CENT.

Baranova AV, Yankovsky NK, Kozlov AP, Lobashev AV, Krukovskaya LL;

WPI; 2003-175241/17.

P-PSDB; ADA84115.

Determining if a nucleic acid is a marker for a phenotype/cell type of interest, by global comparison of expressed sequence tags known to be expressed in the phenotype/cell type with all ESTs expressed in normal tissue.

Claim 23; Page 490-491; 516pp; English.

The invention relates to a novel method for determining if a nucleic acid is a marker for a predetermined phenotype/cell type of interest from a

biological species. The method comprises performing a global comparison of a group of expressed sequence tags (ESTs) known to be expressed in the phenotype/cell type of interest with all ESTs expressed in normal tissue in order to identify ESTs that are preferentially expressed in the phenotype/cell type of interest. A method of the invention is useful for determining whether a nucleic acid is a marker for a predetermined phenotype or cell type of interest from a biological species, preferably Arabidopsis or human. The cell type of interest is an abnormal cell such as a tumour cell, and the predetermined phenotype is a stress-induced phenotype such as hyperosmotic stress or high salt conditions. A method of the invention is also useful for determining the progression of colon cancer in a human, for detecting a tumour cell, and for regulating or preventing the growth of a tumour cell. An antibody of the invention is useful for detecting the absence or presence of peptides encoded by tumour-associated markers. A polypeptide of the invention is useful as an immunogen for vaccinating an animal. The present sequence encodes a tumour-associated antigen of the invention.

Sequence 2430 BP; 430 A; 740 C; 781 G; 479 T; 0 U; 0 Other;

Query Match 100.0%; Score 2405; DB 7; Length 2430;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCACTCTCTCCGAGGTTTCTGCTCTCCGCGCTGTGAGTGTGGGGGCTTGGGTGGGAA 60
DB 8 GCACTCTCTCCGAGGTTTCTGCTCTCCGCGCTGTGAGTGTGGGGGCTTGGGTGGGAA 67
QY 61 GGGGCTGTGCTGAGAGCAGCGCGCTCCCTGGAAGAGAAAGTCTCAGCTGAACGAGCGCG 120
DB 68 GGGGCTGTGCTGAGAGCAGCGCGCTCCCTGGAAGAGAAAGTCTCAGCTGAACGAGCGCG 127
QY 121 CCTAGTTTTCGAGAGGAGATCAGGAGTGTTCGAGCGCTGGAACGAGCGTGC 180
DB 128 CTTAGTTTTCGAGAGGAGATCAGGAGTGTTCGAGCGCTGGAACGAGCGTGC 187
QY 181 GATGAGAGAGCGCGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
DB 188 GATGAGAGAGCGCGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 247
QY 241 GCT 300
DB 248 GCT 307
QY 301 CTTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
DB 308 CTTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 367
QY 361 TGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
DB 368 TGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 427
QY 421 GGGGAGAGCGCGCGCGCGCACTTTCTCATTCAAGGCTGCGCGGCTTTTGTCTATC 480
DB 428 GGGGAGAGCGCGCGCGCGCACTTTCTCATTCAAGGCTGCGCGGCTTTTGTCTATC 487
QY 481 AGAGCGGAGCGCGAGAGTAAAGGGGTGCAAGCGCTTCTCTAGTGTGCTGAGGAG 540
DB 488 AGAGCGGAGCGCGAGAGTAAAGGGGTGCAAGCGCTTCTCTAGTGTGCTGAGGAG 547
QY 541 CTGAGGAGCGCGAGCGCGCGCAAGCGCGAGGAGCGCTGAGAGAGCGAGCGCGAGC 600
DB 548 CTGAGGAGCGCGAGCGCGCGCAAGCGCGAGGAGCGCTGAGAGAGCGAGCGCGAGC 607
QY 601 GCGGAGAGCGAGAGTCAAGCGCGCGCAAGCGCGAGGAGCGCTGAGAGAGCGAGCGAGTGC 660
DB 608 GCGGAGAGCGAGAGTCAAGCGCGCGCAAGCGCGAGGAGCGCTGAGAGAGCGAGCGAGTGC 667
QY 661 GCGGAGAGTGAAGAGCGCGCGCTCTGTCACTGAGAGCACTGTGCGCTGTCTCT 720
DB 668 GCGGAGAGTGAAGAGCGCGCGCTCTGTCACTGAGAGCACTGTGCGCTGTCTCT 727
QY 721 CCCGCTGCGCGCGAGAGTTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780

```
Db 728 CCCCCTGGCCCGAGATTTCTGTGCTGTGTCCGGCTGGCCAAAGCCGGCTCCGAGAC 787
Qy 781 TGGCTTTGTGCCACCGCCCTGTGGCCGGGGCCCTCCGTGCACTGCTCCGAGCTGGG 840
Db 788 TGGCTTTGTGCCACCGCCCTGTGGCCGGGGCCCTCCGTGCACTGCTCCGAGCTGGG 847
Qy 841 CCGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
Db 848 CCGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 907
Qy 901 CCTGAGAGCATGGGGCTCCACTGTGGGCTGAGGCCCAAGAACCCCTGCTGGGAT 960
Db 908 CCTGAGAGCATGGGGCTCCACTGTGGGCTGAGGCCCAAGAACCCCTGCTGGGAT 967
Qy 961 TAGCGATTTGCTGTGTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020
Db 968 TAGCGATTTGCTGTGTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1027
Qy 1021 TTTCCCTCCAGAGCATTAACAGACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
Db 1028 TTTCCCTCCAGAGCATTAACAGACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1087
Qy 1081 CCCCAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140
Db 1088 CCCCAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1147
Qy 1141 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
Db 1148 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1207
Qy 1201 TTTCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260
Db 1208 TTTCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1267
Qy 1261 GTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1320
Db 1268 GTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1327
Qy 1321 CATTTGGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1380
Db 1328 CATTTGGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1387
Qy 1381 TAGGTTCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1440
Db 1388 TAGGTTCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1447
Qy 1441 GGGCTTCCGGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1500
Db 1448 GGGCTTCCGGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1507
Qy 1501 CATCAACTACAGAGACAGCGGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1560
Db 1508 CATCAACTACAGAGACAGCGGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1567
Qy 1561 GTTCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1620
Db 1568 GTTCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1627
Qy 1621 GGGGCACTGTATGGCACTATCCAGGTGAGCCAGGGGCTGTGTGTGTGTGTGTGTGTGT 1680
Db 1628 GGGGCACTGTATGGCACTATCCAGGTGAGCCAGGGGCTGTGTGTGTGTGTGTGTGTGT 1687
Qy 1681 GAGTTCCTCATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1740
Db 1688 GAGTTCCTCATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1747
Qy 1741 GAGTTCCTCATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1800
Db 1748 GAGTTCCTCATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1807
Qy 1801 CCAAGTTTTCTCCGCTTCCATGATGTATGTGAGACCTTTCAGTGTGAAGGGGAGAA 1860
```

```
Db 1808 CCAAGTTTTCTCCGCTTCCATGATGTATGTGAGACCTTTCAGTGTGAAGGGGAGAA 1867
Qy 1861 TGTGGCCACAACCGAGGTGTGAGAGCTTTCAGAGCCCTAGATTTTCTTCAGAGAGTGAA 1920
Db 1868 TGTGGCCACAACCGAGGTGTGAGAGCTTTCAGAGCCCTAGATTTTCTTCAGAGAGTGAA 1927
Qy 1921 CGTCTATGAGTCACTGTGCCAGGSCATGAAAGGCAAGGCTGTGAATGTGCAAGCTTA 1980
Db 1928 CGTCTATGAGTCACTGTGCCAGGSCATGAAAGGCAAGGCTGTGAATGTGCAAGCTTA 1987
Qy 1981 GGGTCCCCCGCCAGCTTTTGGACCTTATGACAGCTTCAACCCAGTGTGTGAAGACTTGGC 2040
Db 1988 GGGTCCCCCGCCAGCTTTTGGACCTTATGACAGCTTCAACCCAGTGTGTGAAGACTTGGC 2047
Qy 2041 ACCTTATGCCCGCCCGATTTCTCAGGCTCCAGAGTCTTTTGGCCACACAGAGACTT 2100
Db 2048 ACCTTATGCCCGCCCGATTTCTCAGGCTCCAGAGTCTTTTGGCCACACAGAGACTT 2107
Qy 2101 CAAAGCAGAAAGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2160
Db 2108 CAAAGCAGAAAGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2167
Qy 2161 ACTGTACGTTCTGTGACAGGCTGTAGTGTCTTACCTGTGCCCTTCACTAGCCCGTACG 2220
Db 2168 ACTGTACGTTCTGTGACAGGCTGTAGTGTCTTACCTGTGCCCTTCACTAGCCCGTACG 2227
Qy 2221 CGCCCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2280
Db 2228 CGCCCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2287
Qy 2281 GGAACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2340
Db 2288 GGAACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2347
Qy 2341 CCAAGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2400
Db 2348 CCAAGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2407
Qy 2401 CCTAC 2405
Db 2408 CCTAC 2412
```

```
RESULT 3
AAD49501
ID AAD49501 standard; cDNA; 2994 BP.
XX
AC AAD49501;
XX
XX 24-MAR-2003 (first entry)
XX
XX Human TRICH-3 cDNA.
XX
XX Human; transporter and ion channel; TRICH; atherosclerosis; cancer;
XX gene therapy; gene; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 50..2485
XX FT /tag= a
XX FT /product= "Human TRICH protein"
XX
XX WO200283712-A2.
XX
XX PN
XX PD 24-OCT-2002.
XX
XX PF 12-APR-2002; 2002WC-US011760.
XX
XX PR 12-APR-2001; 2001US-0283440P.
XX PR 20-APR-2001; 2001US-0285592P.
XX PR 27-APR-2001; 2001US-0287263P.
XX PR 04-MAY-2001; 2001US-0286666P.
```

PR 18-MAY-2001; 2001US-0292042P.
PR 25-MAY-2001; 2001US-0293724P.
PR 22-JAN-2002; 2002US-0351107P.
XX
XX
PA (INCY-) INCYTE GENOMICS INC.
PI Baughn MR, Elliott VS, Hafalia AJA, Yang J, Walla NK, Ramkumar J,
PI Forsythe JD, Lu Y, Tang YT, Yue H, Raumann BE, Lal PG, Azimzai Y,
PI Lu DAM, Gandhi AR, Thornton M, Nguyen DB, Arvizu CS, Emerling BM,
PI Swarnakar AJ, Yao MG, Ding L, He A, Griffin JA, Sanjanwala MM,
PI Gietzen KU, Lee EA, Xu Y, Au-Young JK, Das D, Lee ST, Chang H,
XX WPI; 2003-092996/08.
XX P-PSDB; AAE32069.
PT New human functional transporters and ion channels (TRICH) polypeptides,
PT useful for preparing a composition for diagnosing or treating a disease
PT associated with decreased expression or overexpression of TRICH e.g.
PT cancer.
XX
XX
XX Claim 5; Page 186-187; 204p; English.
XX
XX
XX The invention relates to human transporters and ion channels (TRICH)
XX polypeptides and nucleic acid molecules encoding such polypeptides. TRICH
XX polypeptides are useful for preparing compositions for diagnosing or treating
XX diseases or conditions associated with decreased expression or
XX overexpression of functional TRICH e.g. atherosclerosis or cancer. The
XX invention is useful in gene therapy. The present sequence is human TRICH
XX cDNA
SQ Sequence 2994 BP; 548 A; 876 C; 929 G; 573 T; 0 U; 68 Other:
Query Match 100.0%; Score 2405; DB 7; Length 2994;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACACTCCCTCCGGGTTCTGCTCTCCGCCGCTGAGAGTGGGGGCTGGGTGGGAT 60
DB 235 GACATCTCTCCGGGTTCTGCTCTCCGCCGCTGAGAGTGGGGGCTGGGTGGGAT 294
QY 61 GGGCGTGTGCGACGCGACGCGGCTCCCTGAGAGAGAGTCTCAGCTAGAAGCGGCG 120
DB 295 GGGCGTGTGCGACGCGACGCGGCTCCCTGAGAGAGAGTCTCAGCTAGAAGCGGCG 354
QY 121 CCTAGGTTTTCCGAGAGAGAGATCAGGAGTGTTCGAGCGGCTGGAACGAGCGTCC 180
DB 355 CCTAGGTTTTCCGAGAGAGAGATCAGGAGTGTTCGAGCGGCTGGAACGAGCGTCC 414
QY 181 GATAGAGAGAGGGGCTCCATGAGTCCCTCTGCTGCTGCGCCCTGCTGCTGTCTACG 240
DB 415 GATAGAGAGAGGGGCTCCATGAGTCCCTCTGCTGCTGCGCCCTGCTGCTGTCTACG 474
QY 241 GCTGCTGCTGTGAAGCTACCTCTGCGCGAGTTCGGCTGGCTTCGCGCGACTTGGC 300
DB 475 GCTGCTGCTGTGAAGCTACCTCTGCGCGAGTTCGGCTGGCTTCGCGCGACTTGGC 534
QY 301 CTCTGCGTGGAGAGCTGTGCTGCAAAAGGGCTCTGAGAGTCCGCGCTGGCGCGAGC 360
DB 535 CTCTGCGTGGAGAGCTGTGCTGCAAAAGGGCTCTGAGAGTCCGCGCTGGCGCGAGC 594
QY 361 TGCCTGCGAGCCCGGAAAGTCCCGAGGGGGCTGACGCTGAGCGCTTGGCGGAACT 420
DB 595 TGCCTGCGAGCCCGGAAAGTCCCGAGGGGGCTGACGCTGAGCGCTTGGCGGAACT 654
QY 421 GGGCCGAGCGCGCGCGCGGACACCTTTCTCATTCAGGCTGCGCGCTTTAGTACTC 480
DB 655 GGGCCGAGCGCGCGCGGACACCTTTCTCATTCAGGCTGCGCGCTTTAGTACTC 714
QY 481 AGAGCGAGCGCGAGAGTAAAGAGGCTGACGCGCTTCTTACGTCGCTAGCTGAGCGGA 540
DB 715 AGAGCGAGCGCGAGAGTAAAGAGGCTGACGCGCTTCTTACGTCGCTAGCTGAGCGGA 774
QY 541 CTGGGAGCCCGAGCGCGGACAGCGGCGAGGGAGCGCTGAGAGAGCGGAGCGGCGAGC 600

DB 775 CTGGGAGCCCGAGCGCGGACAGCGGCGAGGGGAGCGCTGAGAGAGCGAGCGGCGAGC 834
QY 601 GCGCGAGCCCGAGAGTACAGCGCGCGGAGGGGAGGTTTCCGAGAGGGGACGGTGC 660
DB 835 GCGCGAGCCCGAGAGTACAGCGCGCGGAGGGGAGGTTTCCGAGAGGGGACGGTGC 894
QY 661 GCGCAGAGTGAAGAGAGCCCGCCCTCTGTGCTGACCTGAGAGCACTGTGGCGCTGCTC 720
DB 895 GCGCAGAGTGAAGAGAGCGCGCGCTCTGTGCTGACCTGAGAGCACTGTGGCGCTGCTC 954
QY 721 CCGCGCTGCGCCAGAGTTTCTGTGCTCTGTGGCTTGGGCTGAGGAGCGCGCTGCGCAC 780
DB 955 CCGCGCTGCGCCAGAGTTTCTGTGCTCTGTGGCTTGGGCTGAGGAGCGCGCTGCGCAC 1014
QY 781 TGCTTTGTGTCACAGCGCCCTGCGCGGGGCGCCCTGTGCTGACCTGCTCGCAGCTGCGG 840
DB 1015 TGCTTTGTGTCACAGCGCCCTGCGCGGGGCGCCCTGTGCTGACCTGCTCGCAGCTGCGG 1074
QY 841 GCGCGCGCGCTGTGCTGCGCGCAAGATTCTTGAAGTCCCTGAGCGCGGACCTGCGCGC 900
DB 1075 GCGCGCGCGCTGTGCTGCGCGCAAGATTCTTGAAGTCCCTGAGCGCGGACCTGCGCGC 1134
QY 901 CCGAGAGCCATGGGGGCTCCAGCTGTGGGCTGAGAGCGCGGAGCCACCTGCTGGAAT 960
DB 1135 CCGAGAGCCATGGGGGCTCCAGCTGTGGGCTGAGAGCGCGGAGCCACCTGCTGGAAT 1194
QY 961 TAGCGATTTCTGCTGATGATGTCGCTGAGTGAAGTGAATGGGCGAGTCCAGAGTACCTTC 1020
DB 1195 TAGCGATTTCTGCTGATGATGTCGCTGAGTGAAGTGAATGGGCGAGTCCAGAGTACCTTC 1254
QY 1021 TTCTCCCGCAGAGCATTAACAACAACGCTGCTGATCACTTTCACTTGGGACCAAGGGGCT 1080
DB 1255 TTCTCCCGCAGAGCATTAACAACAACGCTGCTGATCACTTTCACTTGGGACCAAGGGGCT 1314
QY 1081 CCGCAAGGCTGCTGAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
DB 1315 CCGCAAGGCTGCTGAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1374
QY 1141 GTGTGTGTGTCACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1375 GTGTGTGTGTCACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1434
QY 1201 TTCCCTGCTGGGATGCTGGGCTGATGAGCATTGGGGGCGACAGTGGTGTGAATCAA 1260
DB 1435 TTCCCTGCTGGGATGCTGGGCTGATGAGCATTGGGGGCGACAGTGGTGTGAATCAA 1494
QY 1261 GTTCTCGGCTGTGATGTTCTGGGAAGATTGCCAGACAGAGGATGAGGATGTTCCAGTA 1320
DB 1495 GTTCTCGGCTGTGATGTTCTGGGAAGATTGCCAGACAGAGGATGAGGATGTTCCAGTA 1554
QY 1321 CATTGGGAGACTGTGCGGATACCTGTGCAACGAGCCCGAGAGAGGAGAAAGTGGGCGA 1380
DB 1555 CATTGGGAGACTGTGCGGATACCTGTGCAACGAGCCCGAGAGAGGAGAAAGTGGGCGA 1614
QY 1381 TAAAGTCCGCTGCAAGTGGGCGAGCGGGCTGCGCCAGATACCTGGAGAGCTTTGTGCG 1440
DB 1615 TAAAGTCCGCTGCAAGTGGGCGAGCGGGCTGCGCCAGATACCTGGAGAGCTTTGTGCG 1674
QY 1441 GCGCTTGGGCGCCCTGAGAGTGTGCTGAGAGATATGATGATGATGATGATGATGATGATGAT 1500
DB 1675 GCGCTTGGGCGCCCTGAGAGTGTGCTGAGAGATATGATGATGATGATGATGATGATGATGAT 1734
QY 1501 CATCAACTACAGAGAGAGCGGGGCGGTGTGGGGCGTGTCTTCTGCTTACAGCATAT 1560
DB 1735 CATCAACTACAGAGAGAGCGGGGCGGTGTGGGGCGTGTCTTCTGCTTACAGCATAT 1794
QY 1561 CTTCCTCTTCTCTTGAATGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
DB 1795 CTTCCTCTTCTCTTGAATGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1854
QY 1621 GGGGCACTGTATGGGCAATCTCGAGGTGAGCGAGGGCTGCTGTGGGCGCCCGGTAAGCCA 1680

Db 421 AGCGCGCGCGACACTTTCTCATTTACGGGCTCGCGGCTTTACTACTACAGAGCGG 480
 QY 489 AGCGGAGATTAACAGGGCTGCACGCGCTTCTTACGTGCGTAGAGCTGGAGTGGGAC 548
 Db 481 AGCGGAGATTAACAGGGCTGCACGCGCTTCTTACGTGCGTAGAGCTGGAGTGGGAC 540
 QY 549 CCGAGCGCGCGACAGCGGCGGAGAGCGCTTGAAGAAAGCGAGCGGCGAGCGGAG 608
 Db 541 CCGAGCGCGCGACAGCGGCGGAGAGCGCTTGAAGAAAGCGAGCGGCGAGCGGAG 600
 QY 609 CCGGAGATTCAGAGCGGCGGAGAGCGGAGGTTTCCGAGAGGAGACCGTCCGCAAG 668
 Db 601 CCGAGATTCAGAGCGGCGGAGAGCGGAGGTTTCCGAGAGGAGACCGTCCGCAAG 660
 QY 669 GTGAGAGAGCGCGCGCGCTTGTGACCTGTGAGCAATGTGGCGTGTCTCCCGCTG 728
 Db 661 GTGAGAGAGCGCGCGCGCTTGTGACCTGTGAGCAATGTGGCGTGTCTCCCGCTG 720
 QY 729 GCCGAGATTTCTGTGAGCTGTGTGCGGAGCGGAGGCGGCGGCTGTGCGCATGCGTTTG 788
 Db 721 GCCGAGATTTCTGTGAGCTGTGTGCGGAGCGGAGGCGGCGGCTGTGCGCATGCGTTTG 780
 QY 789 TGCCGAGCGCGCTGCGCGCGGCGCGCTGTGCACTGCGCTCCGAGCTGCGGCGCGG 848
 Db 781 TGCCGAGCGCGCTGCGCGCGGCGCGCTGTGCACTGCGCTCCGAGCTGCGGCGCGG 840
 QY 849 CGCTGTGTGCTGCGCGCGCAAGTTTCTGAGTCCCTTGAAGCGGAGCTGCGCGCGG 908
 Db 841 CGCTGTGTGCTGCGCGCGCAAGTTTCTGAGTCCCTTGAAGCGGAGCTGCGCGCGG 900
 QY 909 CCATGGGGCTCCACTGTGGGCTGAGAGCGGAGCAACCCCTGTGGAATTAAGCAT 968
 Db 901 CCATGGGGCTCCACTGTGGGCTGAGAGCGGAGCAACCCCTGTGGAATTAAGCAT 960
 QY 969 TGCTGTGTGAAGTGTCCGCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1028
 Db 961 TGCTGTGTGAAGTGTCCGCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1020
 QY 1029 AGAGATTAACAGAGAGGCGCTGTACCTTCACTCTGAGAGCAACCGGCGCTCCGCAAG 1088
 Db 1021 AGAGATTAACAGAGAGGCGCTGTACCTTCACTCTGAGAGCAACCGGCGCTCCGCAAG 1080
 QY 1089 CTGCTCGATCATGATCTGAAGATCTGCAATGCGAGGCGTTTATCAGTGTGTGTG 1148
 Db 1081 CTGCTCGATCATGATCTGAAGATCTGCAATGCGAGGCGTTTATCAGTGTGTGTG 1140
 QY 1149 TCCACAGAGAAATGTGATCTTACCTGCGCTTCCCACTTACCAATGTCGCGTCCCTGC 1208
 Db 1141 TCCACAGAGAAATGTGATCTTACCTGCGCTTCCCACTTACCAATGTCGCGTCCCTGC 1200
 QY 1209 TGGGATCGTGGGCTGATGAGGATTTGGGCGCAAGTGTCTGAAATCCAAGTTCTCGG 1268
 Db 1201 TGGGATCGTGGGCTGATGAGGATTTGGGCGCAAGTGTCTGAAATCCAAGTTCTCGG 1260
 QY 1269 CTGCTCAGTCTTGGGAAATGTCAGAGCAAGGCTGATCGGTTTCCAGTATCTGAGG 1328
 Db 1261 CTGCTCAGTCTTGGGAAATGTCAGAGCAAGGCTGATCGGTTTCCAGTATCTGAGG 1320
 QY 1329 AGCTGTGCGATATCTTGTCAACAGAGCGCGGAGCAAGAGTGGCCATTAAGTTC 1388
 Db 1321 AGCTGTGCGATATCTTGTCAACAGAGCGCGGAGCAAGAGTGGCCATTAAGTTC 1380
 QY 1389 GCGTGGCAGTGGGAGCGGCGCTGCGCCAGATACCTGTGGAGCGTTTGTGCGGCGCTTG 1448
 Db 1381 GCGTGGCAGTGGGAGCGGCGCTGCGCCAGATACCTGTGGAGCGTTTGTGCGGCGCTTG 1440
 QY 1449 GCGCGCTGAGGTTGCTGAGAGATTAAGATGAGAGAGGCAAGTGGCCATTAAGTTC 1508
 Db 1441 GCGCGCTGAGGTTGCTGAGAGATTAAGATGAGAGAGGCAAGTGGCCATTAAGTTC 1500
 QY 1509 ACACAGAGAGCGGCGCGCTGTGTGGGCGTGTCTTCTGCTTACAGAGATATCTTCCCT 1568
 Db 1501 ACACAGAGAGCGGCGCGCTGTGTGGGCGTGTCTTCTGCTTACAGAGATATCTTCCCT 1560

QY 1569 TCTCTGTGATGTGCTATGATGTGACCAAGAGAGCAATTCGGGACCCCCAGGGGACT 1628
 Db 1561 TCTCTGTGATGTGCTATGATGATGACCAAGAGAGCAATTCGGGACCCCCAGGGGACT 1620
 QY 1629 GTATGGCCCATCTTCCAGGTGAGCCAGGGGCTGTGGTGGCCCCGGTAAAGCAGAGTCC 1688
 Db 1621 GTATGGCCCATCTTCCAGGTGAGCCAGGGGCTGTGGTGGCCCCGGTAAAGCAGAGTCC 1680
 QY 1689 CATTCGTGGGCTATGCTGCGGCGGCAAGCTGAGCCAGGGGAAATGCTAAAGATGTCT 1748
 Db 1681 CATTCGTGGGCTATGCTGCGGCGGCAAGCTGAGCCAGGGGAAATGCTAAAGATGTCT 1740
 QY 1749 TCGGCTGTGGGATGTTTCTTCAACACTGGGGAAGCTGCTGTGTCTGCAATGACCAAGTT 1808
 Db 1741 TCGGCTGTGGGATGTTTCTTCAACACTGGGGAAGCTGCTGTGTCTGCAATGACCAAGTT 1800
 QY 1809 TTCTCGGCTTCCATGATGCTGAGAGACCTTCAGGTGAGAGGAGGAGAAATGTGGCCA 1868
 Db 1801 TTCTCGGCTTCCATGATGCTGAGAGACCTTCAGGTGAGAGGAGGAGAAATGTGGCCA 1860
 QY 1869 CAACCGAGGTGCGAGAGTCTTGCAGCGCCCTGAGATTTCTTCAAGAGTGAACGTCTAG 1928
 Db 1861 CAACCGAGGTGCGAGAGTCTTGCAGCGCCCTGAGATTTCTTCAAGAGTGAACGTCTAG 1920
 QY 1929 GAGTCACTGTGCGAGGGCAATGAAGGCGAGGCTGGAATGGCAGGCGCTAGTTCTGCGCTCC 1988
 Db 1921 GAGTCACTGTGCGAGGGCAATGAAGGCGAGGCTGGAATGGCAGGCGCTAGTTCTGCGCTCC 1980
 QY 1989 CCCAGCTTTGAGACTTATGAGCTGTACACCCAGCTGTCTGAGAACTTGCAACTTATG 2048
 Db 1981 CCCAGCTTTGAGACTTATGAGCTGTACACCCAGCTGTCTGAGAACTTGCAACTTATG 2040
 QY 2049 CCGGCGCGGATTTCTCAGAGCTTCCAGAGTCTTTGGCCACACAGAGACTTCAACAGC 2108
 Db 2041 CCGGCGCGGATTTCTCAGAGCTTCCAGAGTCTTTGGCCACACAGAGACTTCAACAGC 2100
 QY 2109 AGAAAGTTCGAGTGAAGAAATGAAGGCTTGCAGCCCGAGCACTGTGACCCACTGTAG 2168
 Db 2101 AGAAAGTTCGAGTGAAGAAATGAAGGCTTGCAGCCCGAGCACTGTGACCCACTGTAG 2160
 QY 2169 TTCTGAGACAGGCTGATGCTTACCTGCTTCACTGCTTCAACTGCGGCTACAGGCGCTCC 2228
 Db 2161 TTCTGAGACAGGCTGATGCTTACCTGCTTCACTGCTTCAACTGCGGCTACAGGCGCTCC 2220
 QY 2229 TGGCAGAAACCTTCAATCTGAGAACTTCCACCTGAGGCACTGAGAGAGAACTCT 2288
 Db 2221 TGGCAGAAACCTTCAATCTGAGAACTTCCACCTGAGGCACTGAGAGAGAACTCT 2280
 QY 2289 GTGGGAGTGGGCGCGTGGCAGGTGATCTGAGGCTGTGAGGAGTCTTTCTATACCAAGCT 2348
 Db 2281 GTGGGAGTGGGCGCGTGGCAGGTGATCTGAGGCTGTGAGGAGTCTTTCTATACCAAGCT 2340
 QY 2349 GCGGTCACTATTTTGTAAATGAATGCTGAGAGCTGATCAGAGTGTCTGGA 2400
 Db 2341 GCGGTCACTATTTTGTAAATGAATGCTGAGAGCTGATCAGAGTGTCTGGA 2392

RESULT 5
 AA233977
 ID AA233977 standard; cdna; 2574 bp.
 XX AA233977;
 AC
 XX
 DT 07-DEC-1999 (first entry)
 XX
 DE Human PRO703 nucleotide sequence.
 XX
 XX Human; PNO; EST; expressed sequence tag; PCR primer; hybridisation;
 KM probe; blood coagulation disorder; cancer; cellular adhesion disorder;
 KM secreted protein; transmembrane protein; ss.
 XX
 OS Homo sapiens.

OY	318	IGTGTGCAAAAGGGCTCTTTCGAGCTGCGCGCTGGCCGCGCGCTGGCCGACCCGGAAG	377
Db	374	TGTGTGCAAAAGGGCTCTTTCGAGCTGCGCGCTGGCCGCGCGCTGGCCGACCCGGAAG	433
OY	378	GTCCCGAAGGGGGCTGCAGACCTGGCCCTGGCCCTCGCGGAATTGGCCCGACGCGCG	437
Db	434	GTCCCGAAGGGGGCTGCAGACCTGGCCCTGGCCCTCGCGGAATTGGCCCGACGCGCG	493
OY	438	CGCAACCTTCTCATTTCAAGGCTCGCGGCGCTTAGTACTCAAGGCGGAGCGAGA	497
Db	494	CGCAACCTTCTCATTTCAAGGCTCGCGGCGCTTAGTACTCAAGGCGGAGCGAGA	553
OY	498	GTAACAGGGCTGCAGCGCTTCTTACCTGCGCTAGAGCTGGAGCTGGGACCCGACGCG	557
Db	554	GTAACAGGGCTGCAGCGCTTCTTACCTGCGCTAGAGCTGGAGCTGGGAGCCGACGCG	613
OY	558	GCGACAGCGCGGAGGGGAGCGCTGGAGAAAGCGAGCGGCGCGCGCGGAGCTGGAGT	617
Db	614	GCGACAGCGCGGAGGGGAGCGCTGGAGAAAGCGAGCGGCGCGCGCGGAGCTGGAGT	673
OY	618	CAGCGGCGGGAAGCGGCGGAGTTTGCGGAGGGGAAGCGTGCAGCGGAGGTGAGAG	677
Db	674	CAGCGGCGGGAAGCGGCGGAGTTTGCGGAGGGGAAGCGTGCAGCGGAGGTGAGAG	733
OY	678	CGCGCGCTCTGTACCTGACCTGAGCACTGTGGCGCTGCTCTTCCGCTGGCCGAGAT	737
Db	734	CGCGCGCTCTGTACCTGACCTGAGCACTGTGGCGCTGCTCTTCCGCTGGCCGAGAT	793
OY	738	TTTGTGGCTCTGTTTGGGCTGGCGCAAGGCGCGCTGGGACATGCTTTTGGCCACG	797
Db	794	TTTGTGGCTCTGTTTGGGCTGGCGCAAGGCGCGCTGGGACATGCTTTTGGCCACG	853
OY	798	CCCTGCGCGGGGCGCCCTGTGTGCACTGCTCCGAGCTGGCGCGCGCGCTGTGTC	857
Db	854	CCCTGCGCGGGGCGCCCTGTGTGCACTGCTCCGAGCTGGCGCGCGCGCTGTGTC	913
OY	858	TGGCGCCAGATTTTTGGAGTCCCTGGAGCCGGAACCTTCCCGCTTGAGACCATGGGC	917
Db	914	TGGCGCCAGATTTTTGGAGTCCCTGGAGCCGGAACCTTCCCGCTTGAGACCATGGGC	973
OY	918	TCCACCTTGGGCTGCAGGCCCAAGGAACCAACCTGTGGAATTGCAATTGTGGCTG	977
Db	974	TCCACCTTGGGCTGCAGGCCCAAGGAACCAACCTGTGTGAATTGCAATTGTGGCTG	1033
OY	978	AAGTGTCCGCTGGAAGTGAATGGGCGCAATGCCAGATACCTCTTCCCGCCAGACATPA	1037
Db	1034	AAGTGTCCGCTGGAAGTGAATGGGCGCAATGCCAGATACCTCTTCCCGCCAGACATPA	1093
OY	1038	CAGACAGGCGCTGACATCTTCACTCTMGGACACAGCGGCTCCCGCAAGGCTGTCGA	1097
Db	1094	CAGACAGGCGCTGACATCTTCACTCTMGGACACAGCGGCTCCCGCAAGGCTGTCGA	1153
OY	1098	TCAGTCATCTGAAGTCTGCAATGCCAGGCGTTCTATCAGCTGTGTGTCCACAGG	1157
Db	1154	TCAGTCATCTGAAGTCTGCAATGCCAGGCGTTCTATCAGCTGTGTGTCCACAGG	1213
OY	1158	AAAGTGTATCTACTGCGCCCTCCACACTACCAATATCCGCTCCCGTGGGCAATCG	1217
Db	1214	AAAGTGTATCTACTGCGCCCTCCACACTACCAATATCCGCTCCCGTGGGCAATCG	1273
OY	1218	TGGGCTGATGGGCAATGGGGCCACAGTGTGTCTAAATTCGAAGTTTCGCTGTGCTAGT	1277
Db	1274	TGGGCTGATGGGCAATGGGGCCACAGTGTGTCTAAATTCGAAGTTTCGCTGTGCTAGT	1333
OY	1278	TCTGGGAAGATTGGCAAGCAAGCAGGGTGAAGGTTCAGATCAATTGGGAGCTGTGCC	1337
Db	1334	TCTGGGAAGATTGGCAAGCAAGCAGGGTGAAGGTTCAGATCAATTGGGAGCTGTGCC	1393
OY	1338	GATACCTTGTCAACAGCGCCCGAGCAAGGCGAAGCTGGCCATAGGCTCCGGCTGGAG	1397
Db	1394	GATACCTTGTCAACAGCGCCCGAGCAAGGCGAAGCTGGCCATAGGCTCCGGCTGGAG	1453
OY	1398	TGGGCAAGCGGCTGCGCCCAATACCTGGAGCGTTTGTGTGCGCGCTTCGGGCCCTTGC	1457

Db	1454	TGGCAGAGGGCTGTGGCCACGATACCTCTGGAGGCGTTTGTGTGGCGCTTCGGGCCCTGCG	1513
OY	1458	AGGTGCTGTGAGACACTATGTGACTGACAGAGGGCAACGTGGCCACCATCACTGACACAGAC	1517
Db	1514	AGGTGCTGTGAGACACTATGTGACTGACAGAGGGCAACGTGGCCACCATCACTGACACAGAC	1573
OY	1518	AGCGGGGCGCTGTGGGCGGTGCTTCCTGAGCTTTTACAAGCATATCTCCCTTCCTCTGA	1577
Db	1514	AGCGGGGCGCTGTGGGCGGTGCTTCCTGAGCTTTTACAAGCATATCTCCCTTCCTCTGA	1633
OY	1578	TTCCGTATGATGTACCAACAAGAGAGCCGATTCGGGACCCCCACAGGGGCACTGTATGGCCA	1637
Db	1634	TTCCGTATGATGTACCAACAAGAGAGCCGATTCGGGACCCCCACAGGGGCACTGTATGGCCA	1693
OY	1638	CATCTCCAGGTGAGACCAAGGCGTGTGGTGGGCCCGGTAAGCCACAGTCCCATCTCTCG	1697
Db	1634	CATCTCCAGGTGAGACCAAGGCGTGTGGTGGGCCCGGTAAGCCACAGTCCCATCTCTCG	1753
OY	1698	GCTATGCTGTGGGGCCAGAGCTGTGCCCAAGGGGAAGTTGCTAAGAGATGTCTTCGGGCTG	1757
Db	1754	GCTATGCTGTGGGGCCAGAGCTGTGCCCAAGGGGAAGTTGCTAAGAGATGTCTTCGGGCTG	1813
OY	1758	GGGATGTTTTTTCTTAAACACTGGGGGACCTGTGTGTGTGGATGACCAAGTATTTCCGCT	1817
Db	1814	GGGATGTTTTTTCTTAAACACTGGGGGACCTGTGTGTGTGGATGACCAAGTATTTCTCGCT	1873
OY	1818	TCCATGATCTGACTGTGAGACACCTTTCAGGTGGAAGGGGGAGATGTGTGCCAACAACGAG	1877
Db	1874	TCCATGATCTGACTGTGAGACACCTTTCAGGTGGAAGGGGGAGATGTGTGCCAACAACGAG	1933
OY	1878	TGGCAGAGGCTTTTGAGGCGCTTAAGTATTTCTTACAGAGGTGAAGTCTATGTGAGTCACTG	1937
Db	1934	TGGCAGAGGCTTTTGAGGCGCTTAAGTATTTCTTACAGAGGTGAAGTCTATGTGAGTCACTG	1993
OY	1938	TGCCAGGGCATGAAAGGCAGGGCTGGAATGTGACGCCCTAGTCTCGTCCGCCCAACGCTT	1997
Db	1994	TGCCAGGGCATGAAAGGCAGGGCTGGAATGTGACGCCCTAGTCTCGTCCGCCCAACGCTT	2053
OY	1998	TGAGCTTATGACGCTTACACCAACCAACGTGTCTGAGAACTTGCCACCTTATGCGCGGCCCT	2057
Db	2054	TGAGCTTATGACGCTTACACCAACCAACGTGTCTGAGAACTTGCCACCTTATGCGCGGCCCT	2113
OY	2058	GATTCCCTAGGCTCCAGAGAGTCTTTGGCCACCAACAGAGACCTTCAAAACGAGAGAAAGTTC	2117
Db	2114	GATTCCCTAGGCTCCAGAGAGTCTTTGGCCACCAACAGAGACCTTCAAAACGAGAGAAAGTTC	2173
OY	2118	GGATGGCAATGAGGGCTTCGACCCCAACACCTGTCTGACCCACTGTAGCTTCTGAGCC	2177
Db	2174	GGATGGCAATGAGGGCTTCGACCCCAACACCTGTCTGACCCACTGTAGCTTCTGAGCC	2233
OY	2178	AGGCTGTAGGTGCTTACTGTGCCCTTCACAACTTGCCCGGTAAAGGCCCTCCTGTGAGGAA	2237
Db	2234	AGGCTGTAGGTGCTTACTGTGCCCTTCACAACTTGCCCGGTAAAGGCCCTCCTGTGAGGAA	2293
OY	2238	ACCTTCGATCTGAGAACTTCCACACCTGAGAGGACCTGAGAGAGGAACTCTGTGGGCTGG	2297
Db	2234	ACCTTCGATCTGAGAACTTCCACACCTGAGAGGACCTGAGAGAGGAACTCTGTGGGCTGG	2353
OY	2298	GGGCGCTTGTGAGGTGTACTGGGCTGTCAAGGATCTTTTCTATACAGAACTGCGGTCACT	2357
Db	2354	GGGCGCTTGTGAGGTGTACTGGGCTGTCAAGGATCTTTTCTATACAGAACTGCGGTCACT	2413
OY	2358	ATTTTGTATTAATGTGCTGAGCTGATTCAGAGCTGTCTGAGACTA 2404	
Db	2414	ATTTTGTATTAATGTGCTGAGCTGATTCAGAGCTGTCTGAGACTA 2460	
RESULT 6			
AACT8481			
ID AACT8481 standard; cDNA; 2574 BP.			
AC AACT8481;			

XX 08-FEB-2001 (first entry)
 DT Human PRO703 (UNQ367) nucleotide sequence SEQ ID NO:101.
 DE Human; secreted protein; transmembrane protein; PRO; EST; cytosolic;
 KW expressed sequence tag; detection; cancer; ss.
 XX Homo sapiens.
 OS
 PN MO200053756-A2.
 PD 14-SEP-2000.
 XX 18-FEB-2000; 2000MO-US004341.
 XX 08-MAR-1999; 99MO-US005028.
 PR 12-MAR-1999; 99US-0123957P.
 PR 29-MAR-1999; 99US-0126773P.
 PR 21-APR-1999; 99US-0130232P.
 PR 28-APR-1999; 99US-0131445P.
 PR 14-MAY-1999; 99US-0134287P.
 PR 23-JUN-1999; 99US-0141037P.
 PR 26-JUL-1999; 99US-0145698P.
 PR 29-OCT-1999; 99US-0162506P.
 PR 30-NOV-1999; 99MO-US028313.
 PR 02-DEC-1999; 99MO-US028551.
 PR 16-DEC-1999; 99MO-US030095.
 PR 30-DEC-1999; 99MO-US031243.
 PR 05-JAN-2000; 2000MO-US000219.
 PR 06-JAN-2000; 2000MO-US000277.
 XX 06-JAN-2000; 2000MO-US000376.
 XX (GETH) GENENTECH INC.
 PA Ashkenazi AJ, Baker KP, Botstein D, Desmoyers L, Eaton DL;
 PI Ferrara N, Filvaroff E, Fong S, Gao W, Gerder H, Gerritsen ME;
 PI Goddard AJ, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;
 PI Kijavini IU, Kuo SS, Napier MA, Pan J, Paout NF, Roy MA, Shelton DL;
 PI Stewart TA, Tumas D, Williams PM, Wood WI;
 XX WPI: 2000-611443/58.
 DR P-PDB; AAB44255.
 PT Novel PRO polypeptides and polynucleotides used in detection methods, to
 PT target bioactive molecules to specific cells, and to modulate cellular
 PT activities.
 XX
 PS Claim 2: Fig 38; 636pp; English.
 XX
 CC AAC78458 to AAC78599 represent polynucleotide and EST (expressed sequence
 CC tag) sequences which encode secreted or transmembrane PRO polypeptides.
 CC The PRO polynucleotides and polypeptides have cytostatic activity. The
 CC polynucleotides and polypeptides can be used for detecting the presence
 CC of PRO polypeptides in samples, for linking bioactive molecules to cells
 CC and for modulating biological activities of cells, using the polypeptides
 CC for specific targeting. The polypeptide targeting can be used to kill the
 CC target cells, e.g., for the treatment of cancers. The polypeptide pairs
 CC provide specific targeting of bioactive molecules to cells. AAC78600 to
 CC AAC78987 represent PCR primers and probes used in the isolation of the
 CC PRO polynucleotide sequences
 XX
 SQ Sequence 2574 BP; 470 A; 775 C; 821 G; 508 T; 0 U; 0 Other;
 Query Match 99.3%; Score 2387, DB 3; Length 2574;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 18 CTGCTCTCCGCGCGGTGGAGTGTGGGGGCTGGGTGGGAATGGGCGTGTGCCAGGCA 77
 DB 74 CTGCTCTCCGCGCGGTGGAGTGTGGGGGCTGGGTGGGAATGGGCGTGTGCCAGGCA 133

QY 78 CGCGCGCTCCCTGGAGAGAGAACTCTAGCTAGAAAGACCGCCCTAGGTTTGGAGAG 137
 DB 134 CGCGCGCTCCCTGGAGAGAGAACTCTAGCTAGAAAGAGCGGCTTAGGTTTGGAGAG 193
 QY 138 GAGATATGAGGATGTTTGGCAGCGGCTGGAAACCAACGATGCGATAGAGAGCGGCT 197
 DB 194 GAGATATGAGGATGTTTGGCAGCGGCTGGAAACCAACGATGCGATAGAGAGCGGCT 253
 QY 198 CCATGCTGCTCCCTGCTGCTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 257
 DB 254 CCATGCTGCTCCCTGCTGCTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 313
 QY 258 TACACCTCTGCGCGGAGTTGCGCTGCTCCGCGGAGCTTGAGCTTTGCGTGGAGCTC 317
 DB 314 TACACCTCTGCGCGGAGTTGCGCTGCTCCGCGGAGCTTGAGCTTTGCGTGGAGCTC 373
 QY 318 TGTGCTGCAAAAGGCTCTTTCAGAGCTGCGGCTTGCGGCTGCGGCTGCGGCTGCGG 377
 DB 374 TGTGCTGCAAAAGGCTCTTTCAGAGCTGCGGCTTGCGGCTGCGGCTGCGGCTGCGG 433
 QY 378 GTCCCGAGGGGCTGAGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 437
 DB 434 GTCCCGAGGGGCTGAGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 493
 QY 438 CGACACCTTTCTCATTCACGAGCTGCGGCTTTAGCTACTGAGAGCGAGCGGAG 497
 DB 494 CGACACCTTTCTCATTCACGAGCTGCGGCTTTAGCTACTGAGAGCGAGCGGAGCGGAG 553
 QY 498 GTAACGAGGCTGACGCGCTTCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 557
 DB 554 GTAACGAGGCTGACGCGCTTCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 613
 QY 558 GCGACAGCGGCGAGGAGAGCGCTGAGAGAGCGAGCGGCGAGCGGCGAGCGAGATG 617
 DB 614 GCGACAGCGGCGAGGAGAGCGCTGAGAGAGCGAGCGGCGAGCGGCGAGCGAGATG 673
 QY 618 CACGCGCGGAGAGCGGCGGAGGCTTTCGCGAGGAGGAGCGGCTTTCGCGAGGAGGAG 677
 DB 674 CACGCGCGGAGAGCGGCGGAGGCTTTCGCGAGGAGGAGCGGCTTTCGCGAGGAGGAG 733
 QY 678 CCGCGCGGCTTCTGATCACTGAGCACTGAGCACTGAGCACTGAGCACTGAGCACTGAG 737
 DB 734 CCGCGCGGCTTCTGATCACTGAGCACTGAGCACTGAGCACTGAGCACTGAGCACTGAG 793
 QY 738 TTCTGCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 797
 DB 794 TTCTGCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 853
 QY 798 CCTGCGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 857
 DB 854 CCTGCGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 913
 QY 858 TGGCGCGAAGTTCGAGTTCCTGAGAGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 917
 DB 914 TGGCGCGAAGTTCGAGTTCCTGAGAGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 973
 QY 918 TCCACCTGAGGCTGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 977
 DB 974 TCCACCTGAGGCTGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1033
 QY 978 AAGTGTCCGCTGAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 1037
 DB 1034 AAGTGTCCGCTGAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 1093
 QY 1038 CAGACAGGCTGCTGATCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1097
 DB 1094 CAGACAGGCTGCTGATCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1153
 QY 1098 TCAGTATCTGAAGATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1157
 DB 1154 TCAGTATCTGAAGATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1213

1158 AAGATGATCTACCTGCGCCCTCCCACTCTACCAATGTCGCGTTCCCTGTGGGCATCG 1217
1214 AAGATGATCTACCTGCGCCCTCCCACTCTACCAATGTCGCGTTCCCTGTGGGCATCG 1273
1218 TGCGCTGATGAGGATTTGGGGGCAAGTGTGCTGAATTCAGATTCTCGCTGTGAT 1277
1274 TGCGCTGATGAGGATTTGGGGGCAAGTGTGCTGAATTCAGATTCTCGCTGTGAT 1333
1278 TCTGGGAAGATTGCGAGAGACAGGGGTGAGGTTCAGATTCATTGGGAGCTGTGCC 1337
1334 TCTGGGAAGATTGCGAGAGACAGGGGTGAGGTTCAGATTCATTGGGAGCTGTGCC 1393
1338 GATACCTTGTCAACCAAGCCCCGAGAGAGGAGACGTCGATTAAGTTCGCTGACAG 1397
1394 GATACCTTGTCAACCAAGCCCCGAGAGAGGAGACGTCGATTAAGTTCGCTGACAG 1453
1398 TGCGCAGGGGCTGCGCCCAATACCTGGGAGCGTTTGTGGGGCGCTTGGGGCGCTGCG 1457
1454 TGCGCAGGGGCTGCGCCCAATACCTGGGAGCGTTTGTGGGGCGCTTGGGGCGCTGCG 1513
1458 AGGTGCTGAGACATATGAGCTGACAGAGGGCAACGTGGCCACCATCACTACAGAGAC 1517
1514 AGGTGCTGAGACATATGAGCTGACAGAGGGCAACGTGGCCACCATCACTACAGAGAC 1573
1518 AGCGGGGCGCTGTGGGGGCTGTCTTCTGCTTTCAGACATATCTTCCCTTCTCTGA 1577
1574 AGCGGGGCGCTGTGGGGGCTGTCTTCTGCTTTCAGACATATCTTCCCTTCTCTGA 1633
1578 TTGCGTATGATGTACACACAGAGAGACCAATTCGGGAGCCCCGAGGGGCACTGTATGACCA 1637
1634 TTGCGTATGATGTACACACAGAGAGACCAATTCGGGAGCCCCGAGGGGCACTGTATGACCA 1693
1638 CATCTCCAGGTGAGCAGAGGCGCTGTGTGTCGCCCGGTAAAGCTATCCCATTTCTCTGG 1697
1694 CATCTCCAGGTGAGCAGAGGCGCTGTGTGTCGCCCGGTAAAGCTATCCCATTTCTCTGG 1753
1698 GCTATGCTGCGGGGCAAGCTGCGCCAGGGGAAAGTGTCTAAAGATCTTTCGGGCGCTG 1757
1754 GCTATGCTGCGGGGCAAGCTGCGCCAGGGGAAAGTGTCTAAAGATCTTTCGGGCGCTG 1813
1758 GGGATGTTTCTTCAACACTGGGGACTGTCTGTCTGAGATGACCAAGTTTCTCCGCT 1817
1814 GGGATGTTTCTTCAACACTGGGGACTGTCTGTCTGAGATGACCAAGTTTCTCCGCT 1873
1818 TCCATGATGCTATCTGAGACACCTTCAAGTGTGAAGGGGGAAGATGTGGCCCAACCGAGG 1877
1874 TCCATGATGCTATCTGAGACACCTTCAAGTGTGAAGGGGGAAGATGTGGCCCAACCGAGG 1933
1878 TGCGAGAGGCTTTCAGAGGCGCTAGATTCTTCTCAAGAGGTGAACTGTATGAGTCACTG 1937
1934 TGCGAGAGGCTTTCAGAGGCGCTAGATTCTTCTCAAGAGGTGAACTGTATGAGTCACTG 1993
1938 TGCCAGGGGATGAGAGGCGCTGGAATGCGACCTTATGTTCTGCGTCCCCCGACGCTT 1997
1994 TGCCAGGGGATGAGAGGCGCTGGAATGCGACCTTATGTTCTGCGTCCCCCGACGCTT 2053
1998 TGAGACCTTATGAGGCTCTACCAACGCTGAGAACTGTGCACTTATGAGGCGGCGCC 2057
2054 TGAGACCTTATGAGGCTCTACCAACGCTGAGAACTGTGCACTTATGAGGCGGCGCC 2113
2058 GATTCCTCAGGCTCCAGAGGCTTTCGCGACCAAGACCTTCAAGACGAGAAAGTTTC 2117
2114 GATTCCTCAGGCTCCAGAGGCTTTCGCGACCAAGACCTTCAAGACGAGAAAGTTTC 2173
2118 GATTCCTCAGGCTCCAGAGGCTTTCGCGACCAAGACCTTCAAGACGAGAAAGTTTC 2177
2174 GATTCCTCAGGCTCCAGAGGCTTTCGCGACCAAGACCTTCAAGACGAGAAAGTTTC 2233
2178 AGGCTGTAGGCTCTACCTGCGCTTCAACATGCGCGGATGAGCGCCCTCTCGGAGGAA 2237
2234 AGGCTGTAGGCTCTACCTGCGCTTCAACATGCGCGGATGAGCGCCCTCTCGGAGGAA 2293
2238 ACCTTGAAATCTGAGAACTTCCACACCTGAGGCACTGTGAGAGAACTCTGTGGGGTGG 2297

2294 ACCCTGGAATCTGAGAACTTCCACACTGAGGACCTTGAAGAGAACTCTGTGGGTGG 2353
2298 GGGCCGTTGAGAGTACTGAGCTGTGAGGAGATCTTTCTATACCAAGATGCGGTCACT 2357
2354 GGGCCGTTGAGAGTACTGAGCTGTGAGGAGATCTTTCTATACCAAGATGCGGTCACT 2413
2358 ATTTCGTAATTAATGCTGCTGAGGCTGATCCAGCTGTCTCTGACCTTA 2404
2414 ATTTCGTAATTAATGCTGCTGAGGCTGATCCAGCTGTCTCTGACCTTA 2460
RESULT 7
AAC58239
ID AAC58239 standard; cDNA; 2574 BP.
XX
AC AAC58239;
XX
DT 25-JAN-2001 (first entry)
XX
DE Human PRO703 nucleotide sequence SEQ ID NO:28.
XX
KW Human; tumour; diagnosis; neoplastic disease; identification; cancer;
XX tumorigenesis; detection; neoplastic cell growth; proliferation;
KW cytostatic; antiinflammatory; immunomodulatory; inflammatory disorder;
KW immunological disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200053754-A1.
XX
PD 14-SEP-2000.
XX
PF 06-JAN-2000; 2000WO-US000277.
XX
PR 08-MAR-1999; 99WO-US005028.
XX
PR 12-MAR-1999; 99US-0123957P.
XX
PR 29-MAR-1999; 99US-0126773P.
XX
PR 21-APR-1999; 99US-0130232P.
XX
PR 28-APR-1999; 99US-0131445P.
XX
PR 05-OCT-1999; 99WO-US023089.
XX
PR 30-NOV-1999; 99WO-US028313.
XX
PR 02-DEC-1999; 99WO-US028551.
XX
PR 02-DEC-1999; 99WO-US028564.
XX
PR 30-DEC-1999; 99WO-US031243.
XX
PR 30-DEC-1999; 99WO-US031274.
XX
PA (GENTH) GENENTECH INC.
XX
PI Baker KP, Desauvage FU, Goddard A, Gurney AL, Klein RD, Roy MA;
XX Wood WI;
XX
XX
XX MPI: 2000-572269/53.
XX
XX P-PSDB; AAB24054.
XX
XX
XX New isolated antibody for use in compositions and methods for the
XX diagnosis and treatment of neoplastic cell growth and proliferation in
XX mammals, including humans, and in monitoring tumor treatment.
XX
XX
XX Claim 50; Fig 28; 195pp; English.
XX
XX The present invention describes an isolated antibody (Ab) that binds to
XX one of the human proteins (P) designated PRO213, PRO1350, PRO1449,
XX PRO237, PRO324, PRO324, PRO362, PRO615, PRO531, PRO538, PRO3664, PRO618,
XX PRO772, PRO703, PRO792 or PRO474. The Ab can be used in compositions and
XX methods for the diagnosis and treatment of neoplastic cell growth and
XX proliferation in mammals, including humans. Genes and polypeptides
XX encoded by them, that are amplified in the genome of a tumour cell, can
XX be identified and are useful targets for the treatment and prevention of
XX certain cancers and may be used to monitor tumour treatment. Compounds
XX that inhibit the expression or activity of the identified polypeptides
XX can be identified and used as antagonists. Benign or malignant tumours,
XX inflammatory disorders and immunological disorders can be treated.

CC AAC58123 to AAC58224 represent hybridisation probes and PCR primers used
CC in the isolation of the human PRO sequences. AAC58225 to AAC58241 and
CC AAB24041 to AAB24056 represent human PRO polynucleotide and protein
CC sequences given in the exemplification of the present invention
XX

80 Sequence 2574 BP; 470 A; 775 C; 821 G; 508 T; 0 U; 0 Other;

Query Match 99.3%; Score 2387; DB 3; Length 2574;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 18 CTGCTCTCCGCGCCGCTGTGAGTGTGGGGGCTTGGGTGGGATTTGGGCTGTGTCACGCGA 77
DB |||
QY 74 CTGCTCTCCGCGCCGCTGTGAGTGTGGGGGCTTGGGTGGGATTTGGGCTGTGTCACGCGA 133
DB |||
QY 78 CGCGCGCTCTCTGGAAGAGAGTCTCAGCTAGAAAGAGCGGCTTGGTTCGGAAG 137
DB |||
QY 134 CGCGCGCTCTCTGGAAGAGAGTCTCAGCTAGAAAGAGCGGCTTGGTTCGGAAG 193
DB |||
QY 138 GAGGATCAGGAGATGTTTGGAGAGCGGTGGAAACCAAGCGGTCCGATAGAGAGCGGCT 197
DB |||
QY 194 GAGGATCAGGAGATGTTTGGAGAGCGGTGGAAACCAAGCGGTCCGATAGAGAGCGGCT 253
DB |||
QY 198 CCAATGCTGCTGCTCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 257
DB |||
QY 254 CCAATGCTGCTGCTCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 313
DB |||
QY 258 TACACTCTGGCGCGAGTTGCGTGGCTTCCGAGCACTTGGCTTTTGCGGTGCGAGCTC 317
DB |||
QY 314 TACACTCTGGCGCGAGTTGCGTGGCTTCCGAGCACTTGGCTTTTGCGGTGCGAGCTC 373
DB |||
QY 318 TGGTGTGCAAAAGGGGCTTTCGAGCTGCGCGCTGCGCGCGGCTGCGCGCGCGCGAG 377
DB |||
QY 374 TGTGTGTCAAAAGGGGCTTTCGAGCTGCGCGCTGCGCGCGGCTGCGCGCGCGCGAG 433
DB |||
QY 378 GTCCCGAGGGGGGCTGCAAGCTTGGCTGCGCGCTGCGCGAGCTTGGCGAGAGCGCGCG 437
DB |||
QY 434 GTCCCGAGGGGGGCTGCAAGCTTGGCTGCGCGCTGCGCGAGCTTGGCGAGAGCGCGCG 493
DB |||
QY 438 CGCAGACCTTCTCATTTACAGGCTGCGCGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 497
DB |||
QY 494 CGCAGACCTTCTCATTTACAGGCTGCGCGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 553
DB |||
QY 498 GTAAACAGGCTGCTCAGCGCGCTTCTACGTCGCTAGAGCTGGGACCTGGGAGACCGGAG 557
DB |||
QY 554 GTAAACAGGCTGCTCAGCGCGCTTCTACGTCGCTAGAGCTGGGACCTGGGAGACCGGAG 613
DB |||
QY 558 GCGACAGCGCGCGAGGAGCGCTGGAAGAGCGAGCGCGAGCGCGCGAGCGCGAGAGTG 617
DB |||
QY 614 GCGACAGCGCGCGAGGAGCGCTGGAAGAGCGAGCGCGAGCGCGAGCGCGAGAGTG 673
DB |||
QY 618 CAGCGCGCGAGAGCGCGCGAGGAGGCTTGGCGAGAGCGAGGAGCGCGAGAGTG 677
DB |||
QY 674 CAGCGCGCGAGAGCGCGCGAGGAGGCTTGGCGAGAGCGAGGAGCGCGAGAGTG 733
DB |||
QY 678 CCGCGCGCGCTCTGTCACTGTGAGCACTGTGCGCTGCTCTCTCTCTCTCTCTCTCTCT 737
DB |||
QY 734 CCGCGCGCGCTCTGTCACTGTGAGCACTGTGCGCTGCTCTCTCTCTCTCTCTCTCTCT 793
DB |||
QY 738 TTTCTGTGCTGTGTTTGGGCTGAGCAAGCGCGCTGCGAGAGCGCTTGTGCGCAAG 797
DB |||
QY 794 TTTCTGTGCTGTGTTTGGGCTGAGCAAGCGCGCTGCGAGAGCGCTTGTGCGCAAG 853
DB |||
QY 798 CCTGTGCGCGAGGCGCGCTGTGCACTGTGCGAGCTTGGCGAGCGCGCGCGCTGTGTGC 857
DB |||
QY 854 CCTGTGCGCGAGGCGCGCTGTGCACTGTGCGAGCTTGGCGAGCGCGCGCGCTGTGTGC 913
DB |||
QY 858 TGGCGCGCAGAGTTTCTGAGAGTCCCTGAGAGCGGAGACTTCCGCTTGAAGAGCAATGGGCG 917
DB |||
QY 914 TGGCGCGCAGAGTTTCTGAGAGTCCCTGAGAGCGGAGACTTCCGCTTGAAGAGCAATGGGCG 973
DB |||
QY 918 TCCACTGTGGGCTGAGGCGCGAGAGCACTCTGTGAGATTAAGCAATTTGTGCTGCTG 977
DB |||
```

```
DB 974 TCCACTGTGGGCTGAGGCGCGAGAGCACTCTGTGAGATTAAGCAATTTGTGCTGCTG 1033
QY 978 AAGTGTCCGCTGAGAGTGGAGTGGCGAGTGGCGAGATTAAGCAATTTGTGCTGCTGCTG 1037
DB |||
QY 1034 AAGTGTCCGCTGAGAGTGGAGTGGCGAGTGGCGAGATTAAGCAATTTGTGCTGCTGCTG 1093
DB |||
QY 1038 CAGACAGTGCCTGTACATTTCACTGTGCAACAAGGCGCTTCCCAAGGCTGTCTGGA 1097
DB |||
QY 1094 CAGACAGTGCCTGTACATTTCACTGTGCAACAAGGCGCTTCCCAAGGCTGTCTGGA 1153
DB |||
QY 1098 TCACTGATCTGAAGATCTGCAATTCGAGAGGCTTCAATGAGTGTGCTGCTGCTGCTG 1157
DB |||
QY 1154 TCACTGATCTGAAGATCTGCAATTCGAGAGGCTTCAATGAGTGTGCTGCTGCTGCTG 1213
DB |||
QY 1158 AAGTGTGATCTACCTGCGCTTCCCACTTCAACAATTCGAGTTCCTGTCTGGGATCG 1217
DB |||
QY 1214 AAGTGTGATCTACCTGCGCTTCCCACTTCAACAATTCGAGTTCCTGTCTGGGATCG 1273
DB |||
QY 1218 TGGCTGCAATGGGCTTGGGCGCAAGTGTGCTGAATTCGAAGTTCGAGTGTGCTGCTG 1277
DB |||
QY 1274 TGGCTGCAATGGGCTTGGGCGCAAGTGTGCTGAATTCGAAGTTCGAGTGTGCTGCTG 1333
DB |||
QY 1278 TCTGAGAGATTTGCGCAGCAGCAGCAGGAGTGAAGGTTTCCAGTACATTTGGGAGCTGTGC 1337
DB |||
QY 1334 TCTGAGAGATTTGCGCAGCAGCAGGAGTGAAGGTTTCCAGTACATTTGGGAGCTGTGC 1393
DB |||
QY 1338 GATACCTTGTCAACAGAGCTTCCGAGCAGAGAGAGAGTGTGCGCTGAGCGAG 1397
DB |||
QY 1394 GATACCTTGTCAACAGAGCTTCCGAGCAGAGAGAGAGTGTGCGCTGAGCGAG 1453
DB |||
QY 1398 TGGGCAAGCGGCTGCGCGCAGATACCTGGAGAGGTTTGGGCGGCTTGGGCGGCTTGG 1457
DB |||
QY 1454 TGGGCAAGCGGCTGCGCGCAGATACCTGGAGAGGTTTGGGCGGCTTGGGCGGCTTGG 1513
DB |||
QY 1458 AAGTGTGAGAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 1517
DB |||
QY 1514 AAGTGTGAGAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 1573
DB |||
QY 1518 AAGGCGGCGCTGTGGGCGGCTGCTGCTTCAAGATTAAGATTAAGATTAAGATTAAG 1577
DB |||
QY 1574 AAGGCGGCGCTGTGGGCGGCTGCTGCTTCAAGATTAAGATTAAGATTAAGATTAAG 1633
DB |||
QY 1578 TTCGCTATGATGTCAACACAGAGAGAGCAATTCGAGAGCGCGAGGAGCACTGATAGGCA 1637
DB |||
QY 1634 TTCGCTATGATGTCAACACAGAGAGAGCAATTCGAGAGCGCGAGGAGCACTGATAGGCA 1693
DB |||
QY 1638 CATCTCAAGTGAAGCAGAGGCTGTGTGTGCGCGGTAAGCAGAGTTCCTATTCCTG 1697
DB |||
QY 1694 CATCTCAAGTGAAGCAGAGGCTGTGTGTGCGCGGTAAGCAGAGTTCCTATTCCTG 1753
DB |||
QY 1698 GCTATGCTGGCGGCGCAGAGCTGTGCGCAGAGGAGTGTGTAAGATTCCTTCCGAGCTG 1757
DB |||
QY 1754 GCTATGCTGGCGGCGCAGAGCTGTGCGCAGAGGAGTGTGTAAGATTCCTTCCGAGCTG 1813
DB |||
QY 1758 GGGATGTTTCTTCAACACTGTGGAGACTGTGTGTGCGATGACCAAGTTCCTCTCGCT 1817
DB |||
QY 1814 GGGATGTTTCTTCAACACTGTGGAGACTGTGTGTGCGATGACCAAGTTCCTCTCGCT 1873
DB |||
QY 1818 TCCATGATGTAAGTGAAGCACTTCAAGTGAAGGAGGAGATGAGGCAACAACAGAG 1877
DB |||
QY 1874 TCCATGATGTAAGTGAAGCACTTCAAGTGAAGGAGGAGATGAGGCAACAACAGAG 1933
DB |||
QY 1878 TGGCAGAGGCTTCTGAGGCTTGAATTTCTTCAAGAGTGAAGCTGTATGAGAGTCACTG 1937
DB |||
QY 1934 TGGCAGAGGCTTCTGAGGCTTGAATTTCTTCAAGAGTGAAGCTGTATGAGAGTCACTG 1993
DB |||
QY 1938 TGGCAGAGGATTAAGAGCAGGCTTGAATTTCTTCAAGAGTGAAGCTGTATGAGAGTCACTG 1997
DB |||
QY 1994 TGGCAGAGGATTAAGAGCAGGCTTGAATTTCTTCAAGAGTGAAGCTGTATGAGAGTCACTG 2053
DB |||
QY 1998 TGGACCTTATGAGAGCTTCAACCAAGTGTGAGAACTTGGCACTTATGAGCGCGCGCC 2057
DB |||
QY 2054 TGGACCTTATGAGAGCTTCAACCAAGTGTGAGAACTTGGCACTTATGAGCGCGCGCC 2113
DB |||
```


QY 2058 GATTCTCAGGCTCCAGAGTCTTTGGCCACACAGAGACCTTCAACAGCAAGTTTC 2117
DB 2114 GATTCTCAGGCTCCAGAGTCTTTGGCCACACAGAGACCTTCAACAGCAAGTTTC 2173
QY 2118 GGATGCAAAATGAGGGCTTGACCCCGACGACCTGTCTGACCACTGTACGTTCTGACC 2177
DB 2174 GGATGCAAAATGAGGGCTTGACCCCGACGACCTGTCTGACCACTGTACGTTCTGACC 2233
QY 2178 AGGCTGAGTGGCTTACTGCGCCCTGACAACTGCGCGGTACAGGCGCCCTCTGGCGGAA 2237
DB 2234 AGGCTGAGTGGCTTACTGCGCCCTGACAACTGCGCGGTACAGGCGCCCTCTGGCGGAA 2293
QY 2238 ACCTTCGATCTGAGAACTTCCACACCTGAGGACCTGAGAGAGAACTGTGGGGTGG 2297
DB 2294 ACCTTCGATCTGAGAACTTCCACACCTGAGGACCTGAGAGAGAACTGTGGGGTGG 2353
QY 2298 GGGCGGTGACAGGTGTACTGGGCTGTGAGGATCTTTTCTATACCAAGACTCGGGTACT 2357
DB 2354 GGGCGGTGACAGGTGTACTGGGCTGTGAGGATCTTTTCTATACCAAGACTCGGGTACT 2413
QY 2358 ATTTGTATTAATGTGCTGAGACTGATCCAGCTGTCTGACCTA 2404
DB 2414 ATTTGTATTAATGTGCTGAGACTGATCCAGCTGTCTGACCTA 2460

RESULT 8
ACD42510
ID ACD42510 standard; cDNA, 2574 BP.
XX ACD42510;
AC 09-SEP-2003 (first entry)
DT
XX
DE Novel human secreted and transmembrane protein PRO703 cDNA.
XX
XX Human, secreted and transmembrane protein; PRO; virucide; gene therapy;
KM cell death; growth induction cascade; blood coagulation cascade;
KM viral infection; gene; ss.
XX
OS Homo sapiens.
XX
PN US2003050239-A1.
XX
XX 13-MAR-2003.
PD
XX
PF 15-OCT-2001; 2001US-00978191.
XX

PR 17-OCT-1997; 97US-0062250P.
PR 03-NOV-1997; 97US-0062424P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066364P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077641P.
PR 11-MAR-1998; 98US-0077649P.
PR 12-MAR-1998; 98US-0077791P.
PR 13-MAR-1998; 98US-0078004P.
PR 17-MAR-1998; 98US-0004022P.
PR 17-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078910P.
PR 20-MAR-1998; 98US-0078936P.
PR 20-MAR-1998; 98US-0078939P.
PR 25-MAR-1998; 98US-0079294P.
PR 26-MAR-1998; 98US-0079366P.
PR 27-MAR-1998; 98US-0079663P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079689P.
PR 27-MAR-1998; 98US-0079728P.
PR 30-MAR-1998; 98US-0079820P.
PR 30-MAR-1998; 98US-0079923P.
PR 31-MAR-1998; 98US-0080105P.

PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080165P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080328P.
PR 01-APR-1998; 98US-0080333P.
PR 01-APR-1998; 98US-0080334P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 08-APR-1998; 98US-0081071P.
PR 09-APR-1998; 98US-0081195P.
PR 09-APR-1998; 98US-0081203P.
PR 09-APR-1998; 98US-0081229P.
PR 15-APR-1998; 98US-0081817P.
PR 15-APR-1998; 98US-0081819P.
PR 15-APR-1998; 98US-0081838P.
PR 15-APR-1998; 98US-0081952P.
PR 15-APR-1998; 98US-0081955P.
PR 21-APR-1998; 98US-0082569P.
PR 21-APR-1998; 98US-0082700P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 22-APR-1998; 98US-0082804P.
PR 22-APR-1998; 98US-0082796P.
PR 27-APR-1998; 98US-0083336P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083392P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083558P.
PR 29-APR-1998; 98US-0083559P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083500P.
PR 29-APR-1998; 98US-0083455P.
PR 29-APR-1998; 98US-0083544P.
PR 29-APR-1998; 98US-0083558P.
PR 29-APR-1998; 98US-0083559P.
PR 30-APR-1998; 98US-0083742P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 06-MAY-1998; 98US-0084411P.
PR 06-MAY-1998; 98US-0084498P.
PR 07-MAY-1998; 98US-0084600P.
PR 07-MAY-1998; 98US-0084627P.
PR 07-MAY-1998; 98US-0084637P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 13-MAY-1998; 98US-0085323P.
PR 13-MAY-1998; 98US-0085338P.
PR 13-MAY-1998; 98US-0085339P.
PR 15-MAY-1998; 98US-0085573P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085822P.
PR 15-MAY-1998; 98US-0085889P.
PR 15-MAY-1998; 98US-0085697P.
PR 15-MAY-1998; 98US-0085700P.
PR 15-MAY-1998; 98US-0085704P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086322P.
PR 22-MAY-1998; 98US-0086414P.
PR 22-MAY-1998; 98US-0086440P.
PR 22-MAY-1998; 98US-0086485P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087106P.
PR 28-MAY-1998; 98US-0087208P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0030863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 30-JUL-1998; 98US-0094651P.
PR 11-SEP-1998; 98US-0100038P.
PR 07-OCT-1998; 98US-00169978.

PR 07-OCT-1998; 98WO-US021141.
PR 02-NOV-1998; 98US-00184216.
PR 06-NOV-1998; 98US-00187368.
PR 20-NOV-1998; 98US-0109304P.
PR 07-DEC-1998; 98WO-US024855.
PR 22-DEC-1998; 98US-00202054.
PR 22-DEC-1998; 98US-00218517.
PR 23-DEC-1998; 98US-0113296P.
PR 05-JAN-1999; 98WO-US000106.
PR 05-MAR-1999; 99US-00254465.
PR 08-MAR-1999; 99WO-US005028.
PR 10-MAR-1999; 99US-00265686.
PR 10-MAR-1999; 99WO-US005190.
PR 12-MAR-1999; 99US-00267213.
PR 12-MAR-1999; 99US-0123957P.
PR 29-MAR-1999; 99US-0126773P.
PR 12-APR-1999; 99US-00284291.
PR 21-APR-1999; 99US-0130232P.
PR 26-APR-1999; 99US-0131022P.
PR 28-APR-1999; 99US-0131445P.
PR 14-MAY-1999; 99US-00311833.
PR 14-MAY-1999; 99US-0134287P.
PR 02-JUN-1999; 99WO-US010733.
PR 16-JUN-1999; 99WO-US012252.
PR 23-JUN-1999; 99US-0139557P.
PR 07-JUL-1999; 99US-0142680P.
PR 26-JUL-1999; 99US-0145698P.
PR 28-JUL-1999; 99US-0146222P.
PR 25-AUG-1999; 99US-00380137.
PR 25-AUG-1999; 99US-00380138.
PR 25-AUG-1999; 99US-00380142.
PR 29-OCT-1999; 99US-0162504P.
PR 30-NOV-1999; 99WO-US028313.
PR 02-DEC-1999; 99WO-US028551.
PR 02-DEC-1999; 99WO-US028565.
PR 16-DEC-1999; 99WO-US030095.
PR 30-DEC-1999; 99WO-US031243.
PR 30-DEC-1999; 99WO-US031274.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000277.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 24-FEB-2000; 2000WO-US005004.
PR 02-MAR-2000; 2000WO-US005841.
PR 10-MAR-2000; 2000WO-US006319.
PR 21-MAR-2000; 2000WO-US007352.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014941.
PR 30-MAY-2000; 2000WO-US015264.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000US-00709238.
PR 27-NOV-2000; 2000US-00723749.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001WO-US006520.
PR 22-MAR-2001; 2001US-00816744.
PR 22-MAR-2001; 2001US-00816920.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 25-MAY-2001; 2001WO-US017092.
PR 01-JUN-2001; 2001US-00872035.
PR 01-JUN-2001; 2001WO-US017880.
PR 05-JUN-2001; 2001US-00874503.
PR 14-JUN-2001; 2001US-00882636.
PR 19-JUN-2001; 2001US-00886342.

PR 20-JUN-2001; 2001WO-US019692.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 30-JUL-2001; 2001US-00918585.
XX
XX
XX (GETH) GENENTECH INC.
PI Ashkenazi AJ, Baker KP, Botstein D, Deenoyers L, Eaton DL;
PI Ferrera N, Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME;
Query Match 99.3%; Score 2387; DB 7; Length 2574;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 CTGCTCCGCGCGGTGAGTGTGAGGCGGCTGAGTGAATGAGCGTGTGACAGCA 77
Db 74 CTGCTCCGCGCGGTGAGTGTGAGGCGGCTGAGTGAATGAGCGTGTGACAGCA 133
QY 78 CGGCGGCTCCCTGGAAGGAGAACTCAGCTAGAAAGAGCGGCTTGGATTGGAGG 137
Db 134 CGGCGGCTCCCTGGAAGGAGAACTCAGCTAGAAAGAGCGGCTTGGATTGGAGG 193
QY 138 GAGATCAGGAGATGTTTGAGGCGGCTGGAACGAGCGGTGCGATAGAGAGCGGCT 197
Db 194 GAGATCAGGAGATGTTTGAGGCGGCTGGAACGAGCGGTGCGATAGAGAGCGGCT 253
QY 198 CCAATGCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 257
Db 254 CCAATGCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 313
QY 258 TACACCTGCGCGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 317
Db 314 TACACCTGCGCGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 373
QY 318 TGTGCTGCAAAAGGCTCTTGAAGTCCGCGCTGCGCGCTGCGCGCTGCGCG 377
Db 374 TGTGCTGCAAAAGGCTCTTGAAGTCCGCGCTGCGCGCTGCGCGCTGCGCG 433
QY 378 GTCCGAGGAGGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 437
Db 434 GTCCGAGGAGGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 493
QY 438 CGCACACTTTTCATTCACTGAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 497
Db 494 CGCACACTTTTCATTCACTGAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 553
QY 498 GTTACAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 557
Db 554 GTTACAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 613
QY 558 GCGACAGCGGCGAGGAGCGCTGAGGAGGAGCGAGCGGCGAGCGGCGAGATG 617
Db 614 GCGACAGCGGCGAGGAGCGCTGAGGAGGAGCGAGCGGCGAGCGGCGAGATG 673
QY 618 CAGCGGCGGAGAGCGGCGGAGGTTGCGGAGGAGGAGGTTGCGGAGGAGG 677
Db 674 CAGCGGCGGAGAGCGGCGGAGGTTGCGGAGGAGGAGGTTGCGGAGGAGG 733
QY 678 CGCGGCGGCTGTCACCTGAGGAGCACTGAGGCTGAGGCTGAGGCTGAGG 737
Db 734 CGCGGCGGCTGTCACCTGAGGAGCACTGAGGCTGAGGCTGAGGCTGAGG 793
QY 738 TTCTGTGCTGTGTTGGGCTGAGCGCAAGGCGGCTGAGCGGCTTGTGCGCA 797
Db 794 TTCTGTGCTGTGTTGGGCTGAGCGCAAGGCGGCTTGTGCGCAAGG 853
QY 798 CCTGCGCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 857
Db 854 CCTGCGCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 913
QY 858 TGGCGCCAGAGTTTCTGAGTCCCTGAGCGGAGCTGCGGCTGAGAGCCATGGG 917
Db 914 TGGCGCCAGAGTTTCTGAGTCCCTGAGCGGAGCTGCGGCTTAGAGCCATGGG 973

QY 918 TCACCTGTGAGGCTGAGGAGCCAGAAACCAACCTGCTGGAATTTAGGATTTGCTGCTG 977
 DB 974 TCACCTGTGAGGCTGAGGAGCCAGAAACCAACCTGCTGGAATTTAGGATTTGCTGCTG 1033
 QY 978 AAGGTGCTGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGG 1037
 DB 1034 AAGGTGCTGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGG 1093
 QY 1038 CAGACAGTGTGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGG 1097
 DB 1094 CAGACAGTGTGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGG 1153
 QY 1098 TCAGTGTGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGG 1157
 DB 1154 TCAGTGTGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGG 1213
 QY 1158 AAGATGTGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGG 1217
 DB 1214 AAGATGTGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGG 1273
 QY 1218 TGGGCTGTGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGG 1277
 DB 1274 TGGGCTGTGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGG 1333
 QY 1278 TCTGGGAGGATTTGGCAGAGCAGAGGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGG 1337
 DB 1334 TCTGGGAGGATTTGGCAGAGCAGAGGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGG 1393
 QY 1338 GATACCTTGTACACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1397
 DB 1394 GATACCTTGTACACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1453
 QY 1398 TGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1457
 DB 1454 TGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1513
 QY 1458 AGGTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1517
 DB 1514 AGGTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1573
 QY 1518 AGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1577
 DB 1574 AGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1633
 QY 1578 TTCGCTATGATGTACACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1637
 DB 1634 TTCGCTATGATGTACACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1693
 QY 1638 CATCTCCAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1697
 DB 1694 CATCTCCAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1753
 QY 1698 GGTATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1757
 DB 1754 GGTATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1813
 QY 1758 GGGATGTTTCTTCAACACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1817
 DB 1814 GGGATGTTTCTTCAACACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1873
 QY 1818 TCCATGATGTGATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1877
 DB 1874 TCCATGATGTGATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1933
 QY 1878 TGGCAGAGGCTTTCAGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1937
 DB 1934 TGGCAGAGGCTTTCAGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1993
 QY 1938 TGGCAGAGGCTTTCAGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1997
 DB 1994 TGGCAGAGGCTTTCAGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2053

QY 1998 TGGACCTTATGAGGCTTACACACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2057
 DB 2054 TGGACCTTATGAGGCTTACACACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2113
 QY 2058 GATTCCTCAGGCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2117
 DB 2114 GATTCCTCAGGCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2173
 QY 2118 GGATGGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2177
 DB 2174 GGATGGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2233
 QY 2178 AGGCTGTAGGCTCCTACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2237
 DB 2234 AGGCTGTAGGCTCCTACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2293
 QY 2238 ACCTTCGAACTTGAACACTTCCACACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2297
 DB 2294 ACCTTCGAACTTGAACACTTCCACACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2353
 QY 2298 GGGCCGTTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2357
 DB 2354 GGGCCGTTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2413
 QY 2358 ATTTGTATTAATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2404
 DB 2414 ATTTGTATTAATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2460

RESULT 9
 ACA63545
 ID ACA63545 standard; cDNA; 2574 BP.
 XX ACA63545;
 AC 16-JUN-2003 (first entry)
 XX
 DT
 DE Novel human secreted and transmembrane protein PRO703 cDNA.
 XX
 XX Human; secreted and transmembrane protein; PRO; antiinflammatory;
 XX antiarteriosclerotic; cardiatic; anti-infertility; anti-HIV; cytostatic;
 XX antidiabetic; gene therapy; inflammatory disease; organ failure;
 XX atherosclerosis; cardiac injury; infertility; birth defect;
 XX premature aging; AIDS; cancer; diabetic complication; chromosome mapping;
 XX gene mapping; pharmaceutical; diagnostic; biosensor; bioreactor;
 XX tissue typing; gene; ss.
 XX
 OS Homo sapiens.
 PN US2002192706-A1.
 PD 19-DEC-2002.
 XX
 PF 24-OCT-2001; 2001US-00999832.
 XX
 XX 17-OCT-1997; 97US-0062250P.
 PR 03-NOV-1997; 97US-0064249P.
 PR 13-NOV-1997; 97US-0065311P.
 PR 21-NOV-1997; 97US-0066364P.
 PR 10-MAR-1998; 98US-0077450P.
 PR 11-MAR-1998; 98US-0077632P.
 PR 11-MAR-1998; 98US-0077641P.
 PR 11-MAR-1998; 98US-0077649P.
 PR 12-MAR-1998; 98US-0077791P.
 PR 12-MAR-1998; 98US-0078004P.
 PR 13-MAR-1998; 98US-00040220.
 PR 17-MAR-1998; 98US-0078886P.
 PR 20-MAR-1998; 98US-0078910P.
 PR 20-MAR-1998; 98US-0078936P.
 PR 20-MAR-1998; 98US-0078939P.
 PR 25-MAR-1998; 98US-0079294P.
 PR 26-MAR-1998; 98US-0079656P.
 PR 27-MAR-1998; 98US-0079663P.

PR 27-MAR-1998; 98US-0079664P.
 PR 27-MAR-1998; 98US-0079689P.
 PR 27-MAR-1998; 98US-0079768P.
 PR 27-MAR-1998; 98US-0079786P.
 PR 30-MAR-1998; 98US-0079920P.
 PR 30-MAR-1998; 98US-0079923P.
 PR 31-MAR-1998; 98US-0080105P.
 PR 31-MAR-1998; 98US-0080107P.
 PR 31-MAR-1998; 98US-0080165P.
 PR 01-APR-1998; 98US-0080327P.
 PR 01-APR-1998; 98US-0080328P.
 PR 01-APR-1998; 98US-0080333P.
 PR 01-APR-1998; 98US-0080334P.
 PR 08-APR-1998; 98US-0081049P.
 PR 08-APR-1998; 98US-0081070P.
 PR 08-APR-1998; 98US-0081071P.
 PR 09-APR-1998; 98US-0081195P.
 PR 09-APR-1998; 98US-0081203P.
 PR 09-APR-1998; 98US-0081229P.
 PR 15-APR-1998; 98US-0081817P.
 PR 15-APR-1998; 98US-0081819P.
 PR 15-APR-1998; 98US-0081838P.
 PR 15-APR-1998; 98US-0081952P.
 PR 21-APR-1998; 98US-0081955P.
 PR 21-APR-1998; 98US-0082568P.
 PR 21-APR-1998; 98US-0082569P.
 PR 22-APR-1998; 98US-0082700P.
 PR 22-APR-1998; 98US-0082704P.
 PR 22-APR-1998; 98US-0082797P.
 PR 22-APR-1998; 98US-0082804P.
 PR 23-APR-1998; 98US-0082796P.
 PR 07-OCT-1998; 98WO-US021141.
 PR 20-NOV-1998; 98WO-US024855.
 PR 05-JAN-1999; 99WO-US000106.
 PR 08-MAR-1999; 99WO-US005028.
 PR 10-MAR-1999; 99WO-US005150.
 PR 14-MAY-1999; 99WO-US010733.
 PR 02-JUN-1999; 99WO-US012252.
 PR 30-NOV-1999; 99WO-US028313.
 PR 02-DEC-1999; 99WO-US028551.
 PR 02-DEC-1999; 99WO-US028565.
 PR 16-DEC-1999; 99WO-US030095.
 PR 30-DEC-1999; 99WO-US031243.
 PR 30-DEC-1999; 99WO-US031274.
 PR 05-JAN-2000; 2000WO-US000219.
 PR 06-JAN-2000; 2000WO-US000277.
 PR 11-FEB-2000; 2000WO-US000376.
 PR 11-FEB-2000; 2000WO-US003565.
 PR 18-FEB-2000; 2000WO-US004341.
 PR 24-FEB-2000; 2000WO-US005004.
 PR 02-MAR-2000; 2000WO-US005841.
 PR 10-MAR-2000; 2000WO-US006319.
 PR 21-MAR-2000; 2000WO-US007532.
 PR 30-MAR-2000; 2000WO-US008439.
 PR 17-MAY-2000; 2000WO-US013705.
 PR 22-MAY-2000; 2000WO-US014042.
 PR 30-MAY-2000; 2000WO-US014941.
 PR 02-JUN-2000; 2000WO-US015264.
 PR 28-JUL-2000; 2000WO-US020710.
 PR 24-AUG-2000; 2000WO-US023328.
 PR 01-DEC-2000; 2000WO-US032678.
 PR 20-DEC-2000; 2000WO-US034956.
 PR 28-FEB-2001; 2001WO-US006520.
 PR 22-MAR-2001; 2001WO-US009552.
 PR 25-MAY-2001; 2001WO-US017092.
 PR 01-JUN-2001; 2001WO-US017800.
 PR 20-JUN-2001; 2001WO-US019692.
 PR 29-JUN-2001; 2001WO-US021066.
 PR 09-JUL-2001; 2001WO-US021735.
 PR XX
 PR XX
 PR (GETH) GENENTECH INC.
 PR XX

PI Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;
 PI Ferrara N, Filvaroff E, Fong S, Gerber H, Gertisen ME;
 PI Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ;
 PI Kijavrin IJ, Kuo SS, Napier MA, Pan J, Paoni NF, Roy MA, Shelton DL;
 PI Stewart TA, Tamas D, Williams PM, Wood WI;
 DR WPI, 2003-328860/31.
 DR P-PSDB; AB072207.
 XX
 XX
 PT New secreted and transmembrane nucleic acids and polypeptides, designated
 PT as PRO, useful for treating inflammation, organ failure, atherosclerosis,
 PT cardiac injury, infertility, birth defects, premature aging, AIDS, or
 PT cancer.
 XX
 XX
 PS Claim 2, Fig 38; 453pp, English.
 XX
 CC The invention describes an isolated nucleic acid (i) comprising, or which
 CC is at least 80 % sequence identity to, or the full-length coding sequence
 CC of, any of 118 300-2100 nucleotide sequences, which encodes its
 CC corresponding PRO polypeptide selected from 118 100-700 amino acid
 CC sequences, all given in the specification. The nucleic acids and
 CC polypeptides are useful for treating inflammatory diseases, organ
 CC failure, atherosclerosis, cardiac injury, infertility, birth defects,
 CC premature aging, AIDS, cancer, or diabetic complications. The nucleic
 CC acids are useful as hybridization probes, in chromosome and gene mapping,
 CC and in generating antisense RNA or DNA. The polypeptides are useful as
 CC pharmaceuticals, diagnostics, biosensors or bioreactors. Both are useful
 CC in tissue typing. This sequence encodes a novel human secreted and
 CC transmembrane PRO polypeptide
 XX
 SQ Sequence 2574 BP; 470 A; 775 C; 821 G; 508 T; 0 U; 0 Other;
 SQ
 Query Match 99.3%; Score 2387; DB 7; Length 2574;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 18 CTGCTCCGCCCGGTGTGAGTGTGGGCGCTGTGGAATGGCGTGTCCAGCGCA 77
 DB 74 CTGCTCCGCCCGGTGTGAGTGTGGGCGCTGTGGAATGGCGTGTCCAGCGCA 133
 QY 78 CGGCGCTCCCTGGAAGAGAGAGTCACTAGTGAACGAGCGGCTTAGTTTGGGAAG 137
 DB 134 CGGCGCTCCCTGGAAGAGAGTCACTAGTGAACGAGCGGCTTAGTTTGGGAAG 193
 QY 138 GAGATCAGGAGTGTGGAGCGGCTGTGACAGCGATGCGATAGAGAGAGGAGCT 197
 DB 194 GAGATCAGGAGTGTGGAGCGGCTGTGACAGCGATGCGATAGAGAGAGGAGCT 253
 QY 198 CCAATGCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 257
 DB 254 CCAATGCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 313
 QY 258 TACACTCTGCGCGAGTTGCGCTTCCGAGCGGCTTCCGAGCGGCTTCCGAGCGGCT 317
 DB 314 TACACTCTGCGCGAGTTGCGCTTCCGAGCGGCTTCCGAGCGGCTTCCGAGCGGCT 373
 QY 318 TGTGCTGCAAAAGGCTCTTGCAGCTCGCGCTGCGCGGCTGCGCGGCTGCGGAG 377
 DB 374 TGTGCTGCAAAAGGCTCTTGCAGCTCGCGCTGCGCGGCTGCGCGGCTGCGGAG 433
 QY 378 GTCCGAGGAGGAGGCTGAGCGCTGCGCGCTGCGCGGCTGCGCGGCTGCGCGGCTG 437
 DB 434 GTCCGAGGAGGAGGCTGAGCGCTGCGCGCTGCGCGGCTGCGCGGCTGCGCGGCTG 493
 QY 438 CGCACACCTTTCTCATTCAGCGCTGCGCGCTTCTAGTACTCAGAGCGGAGCGGAGA 497
 DB 494 CGCACACCTTTCTCATTCAGCGCTGCGCGCTTCTAGTACTCAGAGCGGAGCGGAGA 553
 QY 498 GTAACAGGCTGAGCGGCTTCCAGTGTGCTGAGTGTGAGTGTGAGTGTGAGTGTGAG 557
 DB 554 GTAACAGGCTGAGCGGCTTCCAGTGTGCTGAGTGTGAGTGTGAGTGTGAGTGTGAG 613
 QY 558 GCGACAGCGGAGGAGCGCTGAGAGGCTGAGCGGCTGAGCGGCTGAGCGGAGCGGAGT 617

Db 614 GCGAAGCGGAGGAGGCGCTTGAGAAAGCGAGCGGCGAGCGCCGGAGCCCGAGATG 673
Qy 618 CAGCGCGGAGAGCGGCGCGAGATTGCGGAGGAGAGCGATGCCGCAAGGTGAGAG 677
Db 674 CAGCGCGGAGAGCGGCGCGAGATTGCGGAGGAGAGCGATGCCGCAAGGTGAGAG 733
Qy 678 CCGCGCGCGCTCTGTACACTGAGAGAACTGTGTGGCTCTCTCCCGCTGGCCAGAGT 737
Db 734 CCGCGCGCGCTCTGTACACTGAGAGAACTGTGTGGCTCTCTCCCGCTGGCCAGAGT 793
Qy 738 TTCTGTGCTCTGTGTGCGGCTGAGCGAGCGGCGCTGCGACTGCTTTGTGCCACG 797
Db 754 TTCTGTGCTCTGTGTGCGGCTGAGCGAGCGGCGCTGCGACTGCTTTGTGCCACG 853
Qy 758 CCCTGCGCGGCGCGCGCGCTGCTGCACTGCTTCCGAGCTGCGGCGCGCGCTGCTG 857
Db 854 CCCTGCGCGGCGCGCGCGCGCTGCTGCACTGCTTCCGAGCTGCGGCGCGCGCTGCTG 913
Qy 858 TGGCGCGAGATTCTGTGAGTCCCTGAGCGCGGAGCTTCCGCGCGCTGAGAGCACTGGGCG 917
Db 914 TGGCGCGAGATTCTGTGAGTCCCTGAGCGCGGAGCTTCCGCGCGCTGAGAGCACTGGGCG 973
Qy 918 TCCACCTGTGGGCTGAGGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 977
Db 974 TCCACCTGTGGGCTGAGGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1033
Qy 978 AAGTGTGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1037
Db 1034 AAGTGTGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1093
Qy 1038 CAGACAGCTGCTGTACACTTTCACCTGTGAGCAACAGCGGCGCTCCCGAGGCTGTGCGA 1097
Db 1094 CAGACAGCTGCTGTACACTTTCACCTGTGAGCAACAGCGGCGCTCCCGAGGCTGTGCGA 1153
Qy 1098 TCAGTCACTGTGAAGTCTGTGCAATCTGTGCAATCTGTGCAATCTGTGCAATCTGTGCA 1157
Db 1154 TCAGTCACTGTGAAGTCTGTGCAATCTGTGCAATCTGTGCAATCTGTGCAATCTGTGCA 1213
Qy 1158 AAGATGTGATTAACCTGCGCGCTCCCACTTACCAATGCTGCGTCTGCTGCGGAGATG 1217
Db 1214 AAGATGTGATTAACCTGCGCGCTCCCACTTACCAATGCTGCGTCTGCTGCGGAGATG 1273
Qy 1218 TGGGCTGATGAGGCAATTGGGCGCAAGTGTGTGAATTCGAAGTCTGCGGCTGCTGAGT 1277
Db 1274 TGGGCTGATGAGGCAATTGGGCGCAAGTGTGTGAATTCGAAGTCTGCGGCTGCTGAGT 1333
Qy 1278 TCTGGGAGATGTCAGAGCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1337
Db 1334 TCTGGGAGATGTCAGAGCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1393
Qy 1338 GATACCTTGTCAACAGCGCGCGAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1397
Db 1394 GATACCTTGTCAACAGCGCGCGAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1453
Qy 1398 TGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1457
Db 1454 TGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1513
Qy 1458 AAGTGTGAGAGCAATGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1517
Db 1514 AAGTGTGAGAGCAATGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1573
Qy 1518 AAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1577
Db 1574 AAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1633
Qy 1578 TTGCGATGATGTCACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1637
Db 1634 TTGCGATGATGTCACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1693
Qy 1638 CATCTCAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1697
XX |||

Db 1694 CATCTCAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1753
Qy 1698 GCTATGCTGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1757
Db 1754 GCTATGCTGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1813
Qy 1758 GGGATGTTTCTTCAACACTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1817
Db 1814 GGGATGTTTCTTCAACACTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1873
Qy 1818 TCCATGATGCTACTGAGAGCACTTCAAGGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1877
Db 1874 TCCATGATGCTACTGAGAGCACTTCAAGGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1933
Qy 1878 TGGCAGAGGCTTCCGAGGCGCTGAGATTTTCTTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1937
Db 1934 TGGCAGAGGCTTCCGAGGCGCTGAGATTTTCTTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1993
Qy 1938 TGGCAGAGGCTTCCGAGGCGCTGAGATTTTCTTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1997
Db 1994 TGGCAGAGGCTTCCGAGGCGCTGAGATTTTCTTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2053
Qy 1998 TGGACCTTATGAGGCTTCTACACCACTGCTGTGAGAACTTGGCCACTTATGCCCGGCGCC 2057
Db 2054 TGGACCTTATGAGGCTTCTACACCACTGCTGTGAGAACTTGGCCACTTATGCCCGGCGCC 2113
Qy 2058 GATTCTCAGGCTCCAGAGGAGTCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2117
Db 2114 GATTCTCAGGCTCCAGAGGAGTCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2173
Qy 2118 GATGTGAAATGAGGCTTGAACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2177
Db 2174 GATGTGAAATGAGGCTTGAACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2233
Qy 2178 AGGCTGTGAGGCTGCTACCTGCGCGCTCAACAATGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2237
Db 2234 AGGCTGTGAGGCTGCTACCTGCGCGCTCAACAATGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2293
Qy 2238 ACCTTGAATCTGAGAACTTCCACACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2297
Db 2294 ACCTTGAATCTGAGAACTTCCACACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2353
Qy 2298 GGGCGCTTGAAGTGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2357
Db 2354 GGGCGCTTGAAGTGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2413
Qy 2358 ATTTTGAATGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2404
Db 2414 ATTTTGAATGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2460
XX |||

RESULT 10
ACAT1709
ID ACAT1709 standard; cDNA; 2574 BP.
XX
AC ACAT1709;
DT 11-AUG-2003 (first entry)
XX
XX Human secreted and transmembrane polypeptide PRO703 cDNA.
XX Human; ss; gene; thymolytic agent; interferon; interleukin; cytokine;
XX erythropoietin; colony stimulating factor; cancer; colorectal carcinoma;
XX apoptosis related condition; AIDS; amyotrophic lateral sclerosis;
XX inflammatory disease; asthma; atherosclerosis; neurodegenerative disease;
XX gastrointestinal disorder; Alzheimer's disease; Parkinson's disease;
XX hypertension; myocardial ischemia; kidney disease; carcinogenesis;
XX glomerulonephritis; lung disease; pulmonary hypertension; pre-eclampsia;
XX bronchial asthma; gastric ulcer; renal failure; cardiovascular disease;
XX inflammatory bowel disease; reproductive disorder; premature labour.
XX Homo sapiens.
XX

```

PN US2002177553-A1.
XX
XX 28-NOV-2002.
PF 15-OCT-2001; 2001US-00978192.
XX
PR 17-OCT-1997; 97US-0062250P.
PR 03-NOV-1997; 97US-0064249P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0077450P.
PR 10-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077641P.
PR 11-MAR-1998; 98US-0077791P.
PR 12-MAR-1998; 98US-0078004P.
PR 13-MAR-1998; 98US-0080422P.
PR 20-MAR-1998; 98US-0078866P.
PR 20-MAR-1998; 98US-0078910P.
PR 20-MAR-1998; 98US-0078936P.
PR 20-MAR-1998; 98US-0078939P.
PR 25-MAR-1998; 98US-0079294P.
PR 26-MAR-1998; 98US-0079636P.
PR 27-MAR-1998; 98US-0079663P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079689P.
PR 27-MAR-1998; 98US-0079728P.
PR 27-MAR-1998; 98US-0079786P.
PR 30-MAR-1998; 98US-0079920P.
PR 30-MAR-1998; 98US-0079923P.
PR 26-JUN-1998; 98US-00105413.
PR 07-OCT-1998; 98US-00168978.
PR 07-OCT-1998; 98US-0021141.
PR 06-NOV-1998; 98US-00184216.
PR 06-NOV-1998; 98US-00187368.
PR 20-NOV-1998; 98US-0024855.
PR 07-DEC-1998; 98US-00202054.
PR 22-DEC-1998; 98US-00218517.
PR 05-JAN-1999; 99US-0000106.
PR 05-JAN-1999; 99US-00254465.
PR 08-MAR-1999; 99US-00050028.
PR 10-MAR-1999; 99US-00256686.
PR 10-MAR-1999; 99US-0005190.
PR 12-MAR-1999; 99US-00267213.
PR 12-APR-1999; 99US-00284291.
PR 14-MAY-1999; 99US-00311833.
PR 14-MAY-1999; 99US-00310733.
PR 02-JUN-1999; 99US-00312252.
PR 25-AUG-1999; 99US-00380137.
PR 25-AUG-1999; 99US-00380138.
PR 25-AUG-1999; 99US-00380142.
PR 30-NOV-1999; 99US-0038313.
PR 02-DEC-1999; 99US-0038551.
PR 02-DEC-1999; 99US-0038565.
PR 16-DEC-1999; 99US-00300095.
PR 30-DEC-1999; 99US-0031243.
PR 03-DEC-1999; 99US-0031274.
PR 05-JAN-2000; 2000US-00002217.
PR 06-JAN-2000; 2000US-00002277.
PR 11-FEB-2000; 2000US-00003376.
PR 18-FEB-2000; 2000US-00043341.
PR 24-FEB-2000; 2000US-0005004.
PR 02-MAR-2000; 2000US-0005841.
PR 10-MAR-2000; 2000US-00063315.
PR 21-MAR-2000; 2000US-00075332.
PR 30-MAR-2000; 2000US-00084309.
PR 17-MAY-2000; 2000US-0013705.
PR 22-MAY-2000; 2000US-0014042.
PR 30-MAY-2000; 2000US-0014941.
PR 02-JUN-2000; 2000US-0015264.
PR 28-JUL-2000; 2000US-00207710.
PR 24-AUG-2000; 2000US-0023328.

PR 08-NOV-2000; 2000US-00709238.
PR 27-NOV-2000; 2000US-00723749.
PR 01-DEC-2000; 2000US-00832678.
PR 20-DEC-2000; 2000US-00747259.
PR 20-DEC-2000; 2000US-00834956.
PR 28-FEB-2001; 2001US-00806520.
PR 28-FEB-2001; 2001US-00815744.
PR 22-MAR-2001; 2001US-00815920.
PR 22-MAR-2001; 2001US-00809552.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 25-MAY-2001; 2001US-00817092.
PR 01-JUN-2001; 2001US-00872035.
PR 01-JUN-2001; 2001US-00817800.
PR 05-JUN-2001; 2001US-00874503.
PR 14-JUN-2001; 2001US-00882632.
PR 19-JUN-2001; 2001US-00886342.
PR 20-JUN-2001; 2001US-00819692.
PR 29-JUN-2001; 2001US-0021066.
PR 09-JUL-2001; 2001US-0021735.
PR 30-JUL-2001; 2001US-00918585.

XX
XX (SETH ) GENENTECH INC.
XX
XX Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL,
XX Ferrara N, Filvaroff E, Fong S, Gao W, Gerber H, Gertszen ME,
XX Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ,
XX Kijavrin IJ, Kuo SS, Nessler MA, Pan J, Paoni NF, Roy MA, Shelton DL,
XX Stewart TA, Thomas D, Williams PM, Wood WI,
XX
XX WPI; 2003-328499/31.
XX
XX P-PSDB; AB094887.
XX
XX New isolated PRO polypeptides e.g. PRO213, PRO274 and PRO300, for use as
XX pharmaceuticals, diagnostics, biosensors and bioreactors, for identifying
XX modulators of receptor-ligand interactions.
XX
XX
XX Claim 2; SEQ ID NO 101; 55PP; English.
XX
XX The invention relates to an isolated secreted and transmembrane
XX polypeptide, designated as PRO polypeptide. The PRO polypeptide is useful
XX in PRO polypeptide detection methods. The PRO polypeptide is useful for
XX linking a bioactive molecule to a cell. The PRO polypeptide or an
XX antibody against it is useful for modulating a biological activity of a
XX cell. The PRO polypeptide is useful in industrial applications including
XX pharmaceuticals, diagnostics, biosensors and bioreactors. The PRO
XX polypeptide is also useful as a thrombolytic agent, interferon,
XX interleukin, erythropoietin, colony stimulating factor and other
XX cytokines. The PRO polypeptide is useful for treating disease such as
XX cancer e.g. colorectal carcinoma; apoptosis related conditions e.g. AIDS,
XX amyotrophic lateral sclerosis; inflammatory disease e.g. asthma,
XX atherosclerosis; neurodegenerative disease e.g. Alzheimer's disease,
XX Parkinson's disease; cardiovascular disease e.g. hypertension and
XX myocardial ischaemia; kidney disease e.g. renal failure and
XX glomerulonephritis; lung disease e.g. pulmonary hypertension, bronchial
XX asthma; gastrointestinal disorders e.g. gastric ulcer and inflammatory
XX bowel disease; reproductive disorders e.g. premature labour and
XX pre-eclampsia; carcinogenesis. The present sequence represents a cDNA
XX encoding a PRO polypeptide of the invention. Note: The sequence data for
XX this patent did not form part of the printed specification but was
XX obtained in electronic format directly from USPTO at
XX seqdata.uspto.gov/sequence.html?docid=20020177553
XX
XX Sequence 2574 BP; 470 A; 775 C; 821 G; 508 T; 0 U; 0 Other;
SQ
Query Match 99.3%; Score 2387; DB 7; Length 2574;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 18 CTGCTCGGCGCGGTGAGTGGAGGCGCTGGATGGAGTGGCGGTCACAGCGCA 77
Db 74 CTGCTTCGCGCGCGGTGAGTGGAGGCGCTGGATGGAGTGGCGGTCACAGCGCA 133

```


QY	78	GGCGGCGCTCCCTGGAGAGGAAAGTCTGACGTAGAAAGAGCGAGCGCCTAGATTTTTCGGAAG	137
Db	134	CGCGGCGCTCCCTGGAGAGGAAAGTCTGACGTAGAAAGAGCGAGCGCCTAGATTTTTCGGAAG	193
QY	138	GAGGATCAGGAGTGTTTTTCGAGCGGCTTGAAACCAAGCGATGCCGATAGAGGAAGCGGACT	197
Db	194	GAGGATCAGGAGTGTTTTTCGAGCGGCTTGAAACCAAGCGATGCCGATAGAGGAAGCGGACT	253
QY	198	CCATGGCGTCCCTCTGCTGCTGCCCTTCGTGCTGTCTATACCGTGTGCTGTGCTGTAAGC	257
Db	254	CCATGGCGTCCCTCTGCTGCTGCCCTTCGTGCTGTCTATACCGTGTGCTGTGCTGTAAGC	313
QY	258	TACACCTCTGCGCCGACAGTTTGCGCTGGCTTCGCGCGGAACTTGAGCCTTTTCGCTGCGACGTC	317
Db	314	TACACCTCTGCGCCGACAGTTTGCGCTGGCTTCGCGCGGAACTTGAGCCTTTTCGCTGCGACGTC	373
QY	318	TGTGCTGCAAAAAGGCTCTTTGAGCTCGCGCCCTTCGCGCGGCTGCGCCGACCTCGGAG	377
Db	374	TGTGCTGCAAAAAGGCTCTTTGAGCTCGCGCCCTTCGCGCGGCTGCGCCGACCTCGGAG	433
QY	378	GTCCCGAGGGGGGCTGCGACCGTGGCCTTGAGCGCTGCGGAACTGGCCGACGAGCGCGCG	437
Db	434	GTCCCGAGGGGGGCTGCGACCGTGGCCTTGAGCGCTGCGGAACTGGCCGACGAGCGCGCG	493
QY	438	CGACACCTTTCTCATTCACGAGCTTCGCGGCGCTTTAGTACTCAGAGGCGGAGCGCGAGA	497
Db	494	CGACACCTTTCTCATTCACGAGCTTCGCGGCGCTTTAGTACTCAGAGGCGGAGCGCGAGA	553
QY	498	GTAACAGGGCTGCGACGCGCCTTCTCATGCTGTGCTATGAGCTGGGAGCTGGGGAGCCGACGCGC	557
Db	554	GTAACAGGGCTGCGACGCGCCTTCTCATGCTGTGCTATGAGCTGGGAGCTGGGGAGCCGACGCGC	613
QY	558	GCGACAGCGCGAGAGGGAGCCCTGGAGAGAGCGAGCGGACGCGCGCGGAGCCGAGAGATG	617
Db	614	GCGACAGCGCGAGAGGGAGCCCTGGAGAGAGCGAGCGGACGCGCGCGGAGCCGAGAGATG	673
QY	618	CAGCGGCTCGGAAGCGGCGCGGAGTTTGCCTCGAGGGGACGATGCGCCAGAGGTGAGAG	677
Db	674	CAGCGGCTCGGAAGCGGCGCGGAGTTTGCCTCGAGGGGAGCGATGCGCCAGAGGTGAGAG	733
QY	678	CCGCGCGCCCTCTGTACCTGAGAGCAATGATGGCGCTGCTCTCCCGCTGAGCCACAGAT	737
Db	734	CCGCGCGCCCTCTGTACCTGAGAGCAATGATGGCGCTGCTCTCTCCCGCTGAGCCACAGAT	793
QY	738	TTCTGTGGCTGTGTTTCGGGCTTGCCGCAAGCGCGCTTGCGACATGCTTTGTGCCACCG	797
Db	794	TTCTGTGGCTGTGTTTCGGGCTTGCCGCAAGCGCGCTTGCGACATGCTTTGTGCCACCG	853
QY	798	CCCTGGCGCGGGGCCCCCTGCTGACATGCGCTCCGAGGTGGGGCGGGCGCGCTGGTGC	857
Db	854	CCCTGGCGCGGGGCCCCCTGCTGACATGCGCTCCGAGGTGGGGCGGGCGCGCTGGTGC	913
QY	858	TGAGCGCCAGATTTTCTGGAGTCCCTGAGAGCGGAACTTCCCGCGCTTGAGAGCCATGAGGC	917
Db	914	TGAGCGCCAGATTTTCTGGAGTCCCTGAGAGCGGAACTTCCCGCGCTTGAGAGCCATGAGGC	973
QY	918	TTCACCTGTGGGCTGCGAGGCCCGGAACCCACCTGCTGGAATTGCGAATTTTGCTGCGTGC	977
Db	974	TTCACCTGTGGGCTGCGAGGCCCGGAACCCACCTGCTGGAATTGCGAATTTTGCTGCGTGC	1033
QY	978	AAGATGTCGCGTGAAGTGGAGCCAGTGCAGAGTAACTCTCTTCCCGCCGAGAGCATAA	1037
Db	1034	AAGATGTCGCGTGAAGTGGAGCCAGTGCAGAGTAACTCTCTTCCCGCCGAGAGCATAA	1093
QY	1038	CAGACACGTCGCTGTACATTTTACCTTGCGACCAAGGAGCTTCCCAAGGCTGCTCGGA	1097
Db	1094	CAGACACGTCGCTGTACATTTTACCTTGCGACCAAGGAGCTTCCCAAGGCTGCTCGGA	1153
QY	1098	TCAGTCATCTGAAGATCCGTGGAAATGCCAAGGCTTTATCAGCTGTGATGTCCACAGG	1157
Db	1154	TCAGTCATCTGAAGATCCGTGGAAATGCCAAGGCTTTATCAGCTGTGATGTCCACAGG	1213
QY	1158	AAGATGATGTACTCTGCGCCTCCACCTTACACACATGTCCGATTTCCCTGCTGGAGCATG	1217

Db	1214	AAAGTGTGATCTTAACCTCGCCCTCCACCTCTACCAATGTCGGTTCCTGTGGGCAATCG	1273
OY	1218	TGGGCTGGCATGGGCAATTGGGGGCAOAGTGTGCTGAAA	1277
Db	1274	TGGGCTGGCATGGGCAATTGGGGGCAOAGTGTGCTGAAA	1333
OY	1278	TCGGGGAAGATTGGCAGGCAACAAGGTTGAGGTTGTCCAGTAACTTGGGAGCTGTGCC	1337
Db	1334	TCGGGGAAGATTGGCAGGCAACAAGGTTGAGGTTGTCCAGTAACTTGGGAGCTGTGCC	1393
OY	1338	GATACCTGTTCACACGAGCCCGGAGCAAGCAAGTGGCCATTAAGGTCGGGTGGCAG	1397
Db	1394	GATACCTGTTCACACGAGCCCGGAGCAAGCAAGTGGCCATTAAGGTCGGGTGGCAG	1453
OY	1398	TGGGCAAGGGGCTGTGGCCCAAGATACCTGGGAGACCTTTGTGCGGCGCTTGGGGCCCTGCG	1457
Db	1454	TGGGCAAGGGGCTGTGGCCCAAGATACCTGGGAGACCTTTGTGCGGCGCTTGGGGCCCTGCG	1513
OY	1458	AGGTGCTGGAGACATATGGAAGTGAAGAGGCAAGTGGGCAACATCACTACACAGAGAC	1517
Db	1514	AGGTGCTGGAGACATATGGAAGTGAAGAGGCAAGTGGGCAACATCACTACACAGAGAC	1573
OY	1518	AGCGGGGCGCTGTGGGGCGTGTCTTCCTGGCTTTACAGACATATCTTCCCTTCTCTTGA	1577
Db	1574	AGCGGGGCGCTGTGGGGCGTGTCTTCCTGGCTTTACAGACATATCTTCCCTTCTCTTGA	1633
OY	1578	TTGGCTATGATGTCACACAGAGAGGCCAATTGGGAGACCCCAAGGGGCACTGTATGGCA	1637
Db	1634	TTGGCTATGATGTCACACAGAGAGGCCAATTGGGAGACCCCAAGGGGCACTGTATGGCA	1693
OY	1638	CATCTCCAAGTAGAGCCAGGGGCTGTGTGTGCCCGGTAAAGCCAGACATGCCCATTCCTGG	1697
Db	1694	CATCTCCAAGTAGAGCCAGGGGCTGTGTGTGCCCGGTAAAGCCAGACATGCCCATTCCTGG	1753
OY	1698	GCTATGCTGGCGGGGCCAGAGCTGGGCCCAAGGGGAATTTGCTAAAGATCTTTCCGGGCTG	1757
Db	1754	GCTATGCTGGCGGGGCCAGAGCTGGGCCCAAGGGGAATTTGCTAAAGATCTTTCCGGGCTG	1813
OY	1758	GGGATGTTTTCTTCAACACTGGGGAACGTGCTGTGTGAGTAGACCAAGTTCCTCGCT	1817
Db	1814	GGGATGTTTTCTTCAACACTGGGGAACGTGCTGTGTGAGTAGACCAAGTTCCTCGCT	1873
OY	1818	TCCATGATGTTACTTGGAGACACCTTCAAGTGGAGAGGGGAGAAATGTGGCCCAACCCAGG	1877
Db	1874	TCCATGATGTTACTTGGAGACACCTTCAAGTGGAGAGGGGAGAAATGTGGCCCAACCCAGG	1933
OY	1878	TGGCAGAGGTCCTCAGAGGCTCTAAGATTTTTCTCAGAGGTGAAGCTCTATGAGTCACTG	1937
Db	1934	TGGCAGAGGTCCTCAGAGGCTCTAAGATTTTTCTCAGAGGTGAAGCTCTATGAGTCACTG	1993
OY	1938	TGGCAGAGGCAAGAGGCAAGGCTGGAATGCAAGCCCTAATTCTGTGTCCCCCAACGCTT	1997
Db	1994	TGGCAGAGGCAAGAGGCAAGGCTGGAATGCAAGCCCTAATTCTGTGTCCCCCAACGCTT	2053
OY	1998	TGGACCTTATGCAAGCTCTACACCCACGTGTGAAACTTGTGCACACTTATGCCCCGGGCCC	2057
Db	2054	TGGACCTTATGCAAGCTCTACACCCACGTGTGAAACTTGTGCACACTTATGCCCCGGGCCC	2113
OY	2058	GATTTCCTCAGGCTCCAGAGTCTTTGGCCACCACAGAACCTTAAACAGACAGAAAGTTC	2117
Db	2114	GATTTCCTCAGGCTCCAGAGTCTTTGGCCACCACAGAACCTTAAACAGACAGAAAGTTC	2173
OY	2118	GGATGGCAAAATGAGGCTTTCGACCCCAAGCACCTGTCTGACCCACTGTATGCTTTGGAC	2177
Db	2174	GGATGGCAAAATGAGGCTTTCGACCCCAAGCACCTGTCTGACCCACTGTATGCTTTGGAC	2233
OY	2178	AGAGCTGAAAGTGTCTACTGCCCCCTTCAACAATGCCCGGTTACAGAGCCCTCTGTGCAGAA	2237
Db	2234	AGAGCTGAAAGTGTCTACTGCCCCCTTCAACAATGCCCGGTTACAGAGCCCTCTGTGCAGAA	2293
OY	2238	ACCTTGAATCTGAGAACTTCCACACTTGAGGCACTGAGAGAGAACTCTGTGGGCTGG	2297

Db 2294 ACCTTGGAATCTGAGACTTCCACACCTAGGACCTGAGAGAGAACTCTGTGGGCTGG 2353
 QY 2298 GGGCCGCTTCAGAGTGTACTGGGCTGTCAAGGATCTTTTCTATACAGAACTGGGCTCACT 2357
 Db 2354 GGGCCGCTTCAGAGTGTACTGGGCTGTCAAGGATCTTTTCTATACAGAACTGGGCTCACT 2413
 QY 2358 ATTCTTATTAATGTGCTGAGCTGATCCAGCTGCTCTCTCACTCACTA 2404
 Db 2414 ATTCTTATTAATGTGCTGAGCTGATCCAGCTGTCTCTGACCTTA 2460

RESULT 11
 ABX92349
 ID ABX92349 standard; cDNA; 2574 BP.
 AC ABX92349;
 XX
 XX 08-MAY-2003 (first entry)
 XX
 DE cDNA encoding human PRO703 polypeptide.
 XX
 KW Human; PRO polypeptide; secreted and transmembrane protein;
 KW Immune disorder; diabetes; hyper-insulinaemia; hypo-insulinaemia;
 KW cardiac insufficiency; nervous system disorder; kidney disorder;
 KW bone disorder; cartilage disorder; arthritis; tumour; wound healing;
 KW genetic disorder; cytosolic; antidiabetic; antidiabetic;
 KW antidiabetic; anti-tumour; vulnary; antidiabetic; dermatological;
 KW cardiac; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN US2002169284-A1.
 XX
 PD 14-NOV-2002.
 XX
 PF 16-OCT-2001; 2001US-00978697.
 XX
 XX 26-MAY-1981; 81US-00267213.
 PR 17-OCT-1997; 97US-0062250P.
 PR 03-NOV-1997; 97US-0064249P.
 PR 13-NOV-1997; 97US-0065311P.
 PR 21-NOV-1997; 97US-0066364P.
 PR 10-MAR-1998; 98US-0077450P.
 PR 11-MAR-1998; 98US-0077632P.
 PR 11-MAR-1998; 98US-0077641P.
 PR 11-MAR-1998; 98US-0077649P.
 PR 12-MAR-1998; 98US-007791P.
 PR 13-MAR-1998; 98US-0078004P.
 PR 17-MAR-1998; 98US-0004022P.
 PR 20-MAR-1998; 98US-0078886P.
 PR 20-MAR-1998; 98US-0078910P.
 PR 20-MAR-1998; 98US-0078936P.
 PR 25-MAR-1998; 98US-0078939P.
 PR 25-MAR-1998; 98US-0079294P.
 PR 26-MAR-1998; 98US-0079656P.
 PR 27-MAR-1998; 98US-0079663P.
 PR 27-MAR-1998; 98US-0079664P.
 PR 27-MAR-1998; 98US-0079689P.
 PR 27-MAR-1998; 98US-0079728P.
 PR 27-MAR-1998; 98US-0079786P.
 PR 30-MAR-1998; 98US-0079920P.
 PR 30-MAR-1998; 98US-0079923P.
 PR 26-JUN-1998; 98US-0010541P.
 PR 07-OCT-1998; 98US-0016897P.
 PR 07-OCT-1998; 98US-0016897P.
 PR 02-NOV-1998; 98US-0018421P.
 PR 06-NOV-1998; 98US-0018736P.
 PR 20-NOV-1998; 98US-0020248P.
 PR 07-DEC-1998; 98US-0020205P.
 PR 22-DEC-1998; 98US-0021851P.
 PR 05-JAN-1999; 98US-0020010P.
 PR 05-MAR-1999; 98US-0025346P.
 PR 08-MAR-1999; 98US-0025346P.
 PR 08-MAR-1999; 98US-0025346P.

PR 10-MAR-1999; 99US-00265686.
 PR 10-MAR-1999; 99US-00265686.
 PR 12-APR-1999; 99US-00284291.
 PR 14-MAY-1999; 99US-00311832.
 PR 14-MAY-1999; 99US-00311832.
 PR 02-JUN-1999; 99US-00311832.
 PR 25-AUG-1999; 99US-00380137.
 PR 25-AUG-1999; 99US-00380137.
 PR 25-AUG-1999; 99US-00380137.
 PR 25-AUG-1999; 99US-00380137.
 PR 30-NOV-1999; 99US-00380137.
 PR 30-NOV-1999; 99US-00380137.
 PR 02-DEC-1999; 99US-00380137.
 PR 16-DEC-1999; 99US-00380137.
 PR 30-DEC-1999; 99US-00380137.
 PR 05-JAN-2000; 99US-00380137.
 PR 06-JAN-2000; 99US-00380137.
 PR 06-JAN-2000; 99US-00380137.
 PR 11-FEB-2000; 99US-00380137.
 PR 18-FEB-2000; 99US-00380137.
 PR 24-FEB-2000; 99US-00380137.
 PR 02-MAR-2000; 99US-00380137.
 PR 10-MAR-2000; 99US-00380137.
 PR 21-MAR-2000; 99US-00380137.
 PR 30-MAR-2000; 99US-00380137.
 PR 17-MAY-2000; 99US-00380137.
 PR 22-MAY-2000; 99US-00380137.
 PR 30-MAY-2000; 99US-00380137.
 PR 02-JUN-2000; 99US-00380137.
 PR 28-JUL-2000; 99US-00380137.
 PR 28-JUL-2000; 99US-00380137.
 PR 28-AUG-2000; 99US-00380137.
 PR 08-NOV-2000; 99US-00380137.
 PR 27-NOV-2000; 99US-00380137.
 PR 01-DEC-2000; 99US-00380137.
 PR 20-DEC-2000; 99US-00380137.
 PR 28-FEB-2001; 99US-00380137.
 PR 28-FEB-2001; 99US-00380137.
 PR 22-MAR-2001; 99US-00380137.
 PR 22-MAR-2001; 99US-00380137.
 PR 22-MAR-2001; 99US-00380137.
 PR 10-MAY-2001; 99US-00380137.
 PR 10-MAY-2001; 99US-00380137.
 PR 25-MAY-2001; 99US-00380137.
 PR 01-JUN-2001; 99US-00380137.
 PR 01-JUN-2001; 99US-00380137.
 PR 05-JUN-2001; 99US-00380137.
 PR 14-JUN-2001; 99US-00380137.
 PR 19-JUN-2001; 99US-00380137.
 PR 20-JUN-2001; 99US-00380137.
 PR 29-JUN-2001; 99US-00380137.
 PR 09-JUL-2001; 99US-00380137.
 PR 30-JUL-2001; 99US-00380137.

(GENTH) GENENTECH INC.
 PA
 XX
 PI Ashkenazi A, Baker KP, Borstein D, Desnoyers L, Eaton D;
 PI Ferrara N, Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME;
 PI Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan KD; Shelton DL;
 PI Kijavlin IJ, Kuo SS, Napier MA, Pan J, Paoni NF, Roy MA, Stewart TA;
 PI Stewart TA, Tumas D, Williams PM, Wood WI;
 XX
 XX WPI: 2003-288163/28.
 DR P-PSDB; AB01085.
 DR
 XX
 XX Novel secreted and transmembrane polypeptides and polynucleotides
 PT encoding them useful for treating cancer, kidney diseases, bone,
 PT cartilage disorders and immune deficiencies.
 XX
 PS Claim 2; Fig 38; 459pp; English.
 XX
 CC The present invention relates to the isolation of novel human PRO
 CC polypeptides, and the polynucleotide sequences encoding them. The PRO
 CC polypeptides are secreted and transmembrane proteins. The PRO


```
DB 1874 TCCATGATCGTACTGAGACACCTTCAGGTGGAAGGGGAGATGTGGCCACAAACGAGG 1933
QY 1878 TGGCAGAGGTCTTCAGAGCCCTAGATTTCTTCAGAGAGTGAACGTCTATGAGTCACTG 1937
DB 1934 TGGCAGAGGTCTTCAGAGCCCTAGATTTCTTCAGAGAGTGAACGTCTATGAGTCACTG 1933
QY 1938 TGGCAGAGGTCTTCAGAGCCCTAGATTTCTTCAGAGAGTGAACGTCTATGAGTCACTG 1937
DB 1994 TGGCAGAGGTCTTCAGAGCCCTAGATTTCTTCAGAGAGTGAACGTCTATGAGTCACTG 2053
QY 1998 TGGCAGAGGTCTTCAGAGCCCTAGATTTCTTCAGAGAGTGAACGTCTATGAGTCACTG 2057
DB 2054 TGGCAGAGGTCTTCAGAGCCCTAGATTTCTTCAGAGAGTGAACGTCTATGAGTCACTG 2113
QY 2058 GATTCCTGAGGCTCCAGAGGCTTTGGCCACCAAGAGACCTTCAAGAGAGAAAGTTC 2117
DB 2114 GATTCCTGAGGCTCCAGAGGCTTTGGCCACCAAGAGACCTTCAAGAGAGAAAGTTC 2173
QY 2118 GATTCCTGAGGCTCCAGAGGCTTTGGCCACCAAGAGACCTTCAAGAGAGAAAGTTC 2177
DB 2174 GATTCCTGAGGCTCCAGAGGCTTTGGCCACCAAGAGACCTTCAAGAGAGAAAGTTC 2233
QY 2178 AGGCTGATAGTGCCTACCTGCCCCCTCACAACCTGCCCCGTACAGCGCCCTCTGACAGGAA 2237
DB 2234 AGGCTGATAGTGCCTACCTGCCCCCTCACAACCTGCCCCGTACAGCGCCCTCTGACAGGAA 2293
QY 2238 AGCTTCGATCTGAGAACTTCACACCTGAGGACCTGAGAGAGGAACTCTGTGGGGTGG 2297
DB 2294 AGCTTCGATCTGAGAACTTCACACCTGAGGACCTGAGAGAGGAACTCTGTGGGGTGG 2353
QY 2298 GGGCGGTGACAGGTGCTGAGGCTGCTGAGGAGTCTTTCTATACCAAGCTGCGGTCACT 2357
DB 2354 GGGCGGTGACAGGTGCTGAGGCTGCTGAGGAGTCTTTCTATACCAAGCTGCGGTCACT 2413
QY 2358 ATTTTGTATTAATGCTGCTGAGGCTGCTGAGGAGTCTTTCTATACCAAGCTGCGGTCACT 2404
DB 2414 ATTTTGTATTAATGCTGCTGAGGCTGCTGAGGAGTCTTTCTATACCAAGCTGCGGTCACT 2460

RESULT 12
ID ACA66090 standard; cDNA; 2574 BP.
AC ACA66090;
AC 24-JUN-2003 (first entry)
DT XX
DE Human cDNA encoding secreted/transmembrane protein PRO703.
XX
XX Human; ss; gene; secreted protein; transmembrane protein; PRO;
XX malignancy; cancer; ovarian cancer; colorectal cancer; sarcoma;
XX leukemia; lymphoma; inflammatory disease; necrosis; atherosclerosis;
XX infertility; premature aging; psoriasis; inflammatory disease;
XX renal disease; arthritis; immune-mediated alopecia; stroke; encephalitis;
XX hepatitis; multiple sclerosis; gene therapy.
OS Homo sapiens.
PN US2003004102-A1.
XX
PD 02-JAN-2003.
XX
XX 15-OCT-2001; 2001US-00978189.
XX
XX 17-OCT-1997; 97US-0062250P.
XX 03-NOV-1997; 97US-0064249P.
XX 13-NOV-1997; 97US-0065311P.
XX 21-NOV-1997; 97US-0066364P.
XX 10-MAR-1998; 98US-0077450P.
XX 11-MAR-1998; 98US-0077632P.
XX 11-MAR-1998; 98US-0077641P.
XX 11-MAR-1998; 98US-0077649P.
XX 12-MAR-1998; 98US-0077791P.
PR 13-MAR-1998; 98US-0078004P.
PR 17-MAR-1998; 98US-0040222P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078910P.
PR 20-MAR-1998; 98US-0078936P.
PR 20-MAR-1998; 98US-0078939P.
PR 25-MAR-1998; 98US-0079294P.
PR 25-MAR-1998; 98US-0079656P.
PR 27-MAR-1998; 98US-0079663P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079689P.
PR 27-MAR-1998; 98US-0079728P.
PR 27-MAR-1998; 98US-0079786P.
PR 30-MAR-1998; 98US-0079920P.
PR 30-MAR-1998; 98US-0079923P.
PR 26-JUN-1998; 98US-0010541P.
PR 07-OCT-1998; 98US-0016897P.
PR 07-OCT-1998; 98US-0021141P.
PR 02-NOV-1998; 98US-0018421P.
PR 06-NOV-1998; 98US-0018736P.
PR 20-NOV-1998; 98US-0024855P.
PR 07-DEC-1998; 98US-0020054P.
PR 22-DEC-1998; 98US-0021851P.
PR 05-JAN-1999; 99US-0000106P.
PR 05-MAR-1999; 99US-0025446P.
PR 08-MAR-1999; 99US-0005028P.
PR 10-MAR-1999; 99US-0026568P.
PR 10-MAR-1999; 99US-0005190P.
PR 12-MAR-1999; 99US-0026721P.
PR 12-MAR-1999; 99US-0028429P.
PR 14-MAY-1999; 99US-0031832P.
PR 14-MAY-1999; 99US-0031832P.
PR 02-JUN-1999; 99US-0010733P.
PR 25-AUG-1999; 99US-0038013P.
PR 25-AUG-1999; 99US-0038013P.
PR 25-AUG-1999; 99US-0038013P.
PR 25-AUG-1999; 99US-0038013P.
PR 30-NOV-1999; 99US-0028313P.
PR 02-DEC-1999; 99US-0028313P.
PR 02-DEC-1999; 99US-0028313P.
PR 16-DEC-1999; 99US-0028313P.
PR 30-DEC-1999; 99US-0031243P.
PR 30-DEC-1999; 99US-0031243P.
PR 05-JAN-2000; 2000US-0000219P.
PR 05-JAN-2000; 2000US-0000219P.
PR 06-JAN-2000; 2000US-0000376P.
PR 11-FEB-2000; 2000US-0000365P.
PR 18-FEB-2000; 2000US-0000365P.
PR 24-FEB-2000; 2000US-0000504P.
PR 01-MAR-2000; 2000US-0000560P.
PR 02-MAR-2000; 2000US-0000584P.
PR 10-MAR-2000; 2000US-0000584P.
PR 21-MAR-2000; 2000US-0000753P.
PR 30-MAR-2000; 2000US-0000843P.
PR 17-MAY-2000; 2000US-0013705P.
PR 22-MAY-2000; 2000US-0014042P.
PR 30-MAY-2000; 2000US-0014941P.
PR 02-JUN-2000; 2000US-0015264P.
PR 28-JUN-2000; 2000US-0020271P.
PR 24-AUG-2000; 2000US-0023328P.
PR 08-NOV-2000; 2000US-00709238P.
PR 10-NOV-2000; 2000US-0030873P.
PR 27-NOV-2000; 2000US-00723749P.
PR 01-DEC-2000; 2000US-0032678P.
PR 20-DEC-2000; 2000US-00747259P.
PR 20-DEC-2000; 2000US-0034956P.
PR 28-FEB-2001; 2001US-0006520P.
PR 22-MAR-2001; 2001US-00816920P.
PR 22-MAR-2001; 2001US-00816920P.
PR 10-MAY-2001; 2001US-00854208P.
PR 10-MAY-2001; 2001US-00854280P.
PR 25-MAY-2001; 2001US-00854280P.
PR 01-JUN-2001; 2001US-00872035P.
```

PR 01-JUN-2001; 2001WO-US017800.
 PR 05-JUN-2001; 2001US-00874503.
 PR 14-JUN-2001; 2001US-00882636.
 PR 19-JUN-2001; 2001US-00886342.
 PR 20-JUN-2001; 2001WO-US019692.
 PR 29-JUN-2001; 2001WO-US021066.
 PR 09-JUL-2001; 2001WO-US021735.
 PR 30-JUL-2001; 2001US-00918585.
 XX (GENE) GENENTECH INC.
 XX Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL,
 PI Ferrara N, Flivarcit E, Fong S, Gao W, Gerber H, Gerritsen ME,
 PI Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ,
 PI Kljavin IJ, Kuo SS, Napier MA, Pan J, Paoni NF, Roy MA, Shelton DL,
 PI Stewart TA, Tumas D, Williams PM, Wood WJ;
 DR WPI, 2003-341189/32.
 DR P-PsDB; AB080354.
 XX New genes and secreted and transmembrane polypeptides (e.g. PRO337 or
 PT PRO159), useful for treating or diagnosing e.g. cancers,
 PT atherosclerosis, infertility, stroke, encephalitis, hepatitis or multiple
 PT sclerosis in mammals.
 XX Claim 2, Fig 38; 460pp; English.
 XX The invention relates to a new isolated nucleic acid molecule comprising a
 CC sequence with at least 80% identity to: (a) a nucleotide encoding any of
 CC 94 PRO polypeptides whose sequences are fully defined in the
 CC specification; or (b) any of 94 nucleotide sequences fully defined in the
 CC specification; or the full length coding sequence of any these 94
 CC nucleotide sequences. Also included are an isolated PRO polypeptide
 CC scoring at least 80% positives when compared to any of the PRO
 CC polypeptide sequences cited above (or an isolated PRO polypeptide having
 CC at least 80% amino acid sequence identity to: (a) an amino acid sequence
 CC encoded by the nucleotide deposited with ATCC numbers listed in the
 CC specification; (b) the PRO polypeptide, lacking its associated signal
 CC peptide; or (c) an extracellular domain of the PRO polypeptide, with or
 CC lacking its associated signal peptide), a vector comprising the nucleic
 CC acid molecule, a host cell comprising the vector (and producing a PRO
 CC polypeptide), a chimeric molecule comprising the PRO polypeptide fused
 CC to a heterologous amino acid sequence and an anti-PRO antibody. The PRO
 CC polypeptides or polynucleotides are useful as pharmaceuticals,
 CC diagnostics, biosensors or bioreactors. These are particularly useful for
 CC detecting or treating e.g. malignancies or cancers (e.g. ovarian cancer,
 CC colorectal cancer, sarcoma, leukemia or lymphoma), inflammatory disease,
 CC necrosis, atherosclerosis, infertility, premature aging, psoriasis,
 CC inflammatory disease, renal disease, arthritis, immune-mediated alopecia,
 CC stroke, encephalitis, hepatitis, or multiple sclerosis in mammals. The
 CC PRO polypeptides are useful in drug screening, particularly as targets
 CC for therapeutic intervention in these diseases, and in the diagnostic
 CC determination of the presence of these diseases. The PRO polypeptides are
 CC also useful as molecular weight markers, or for chromosome
 CC identification. The PRO genes are useful as hybridisation probes, or for
 CC screening libraries of human cDNA, genomic DNA or mRNA. The PRO genes may
 CC also be used in gene therapy, particularly for replacing a defective
 CC gene. The present sequence encodes a PRO polypeptide
 XX
 SQ Sequence 2574 BP; 470 A; 775 C; 821 G; 508 T; 0 U; 0 Other;
 Query Match 99.3%; Score 2387; DB 7; Length 2574;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 138 GAGATCAGGATGTTTGGAGCGGCTGGAACAGACGGTGCATAGAGGAAGCGGCT 197
 Db 194 GAGATCAGGATGTTTGGAGCGGCTGGAACAGACGGTGCATAGAGGAAGCGGCT 253
 QY 198 CCAATGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 257
 Db 254 CCAATGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 313
 QY 258 TACACCTGAGCGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 317
 Db 314 TACACCTGAGCGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 373
 QY 318 TGTGCTGAAAGAGGCTTTCAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 377
 Db 374 TGTGCTGAAAGAGGCTTTCAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 433
 QY 378 GTCCGAGGAGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 437
 Db 434 GTCCGAGGAGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 493
 QY 438 CGCACACTTTCATTCACGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 497
 Db 494 CGCACACTTTCATTCACGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 553
 QY 498 GTACAGGAGCTGACGCGCTTTCAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 557
 Db 554 GTACAGGAGCTGACGCGCTTTCAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 613
 QY 558 GGCACAGCGGCGAGGAGGAGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 617
 Db 614 GGCACAGCGGCGAGGAGGAGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 673
 QY 618 CAGCGGCGGAGAGCGGCGGAGTTTCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 677
 Db 674 CAGCGGCGGAGAGCGGCGGAGTTTCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 733
 QY 678 CGCGCGCGCTTGTACCTGAGCAACTGTGGCGCTGCTTCCCGCGCGCGAGAGT 737
 Db 734 CGCGCGCGCTTGTACCTGAGCAACTGTGGCGCTGCTTCCCGCGCGAGAGT 793
 QY 738 TTCTGTGCTTGTGTGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 797
 Db 794 TTCTGTGCTTGTGTGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 853
 QY 798 CCTGTGCGGCGGCGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 857
 Db 854 CCTGTGCGGCGGCGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 913
 QY 858 TGCGCGCAGAGTTTGTGAGTCCCTGAGACCGGAGCTGCGGCTGAGAGCGATGGGC 917
 Db 914 TGCGCGCAGAGTTTGTGAGTCCCTGAGACCGGAGCTGCGGCTGAGAGCGATGGGC 973
 QY 918 TCCACTGTGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 977
 Db 974 TCCACTGTGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1033
 QY 978 AAGTGTCCGTTGAGTGGAGTGGGCTGAGTGGGCTGAGTGGGCTGAGTGGGCTGAGT 1037
 Db 1034 AAGTGTCCGTTGAGTGGAGTGGGCTGAGTGGGCTGAGTGGGCTGAGTGGGCTGAGT 1093
 QY 1038 CAGACAGTGTGCTGATCATCTTCACTTGTGAGCAAGGAGCTGCCAGAGGCTGTGCA 1097
 Db 1094 CAGACAGTGTGCTGATCATCTTCACTTGTGAGCAAGGAGCTGCCAGAGGCTGTGCA 1153
 QY 1098 TCAATCATGAGAGTCTGAGATGCGAGGAGCTTATGAGCTGTGAGTGTGATGATGAT 1157
 Db 1154 TCAATCATGAGAGTCTGAGATGCGAGGAGCTTATGAGCTGTGAGTGTGATGATGAT 1213
 QY 1158 AAGATGATCTACTGCGGCTTCCACTGACATGTCGAGTTCCTGCTGGGATGCTG 1217
 Db 1214 AAGATGATCTACTGCGGCTTCCACTGACATGTCGAGTTCCTGCTGGGATGCTG 1273
 QY 1218 TGGGCTGATGAGGATTTGGGCGCAGTGTGTGTAATTCAGTTCCTGGGCTGTCACT 1277

Db 1274 TGGGCTGATGGGCAATGGGGCCACAGTGTGCTGATAATCCAAAGTTCTGGGTGGTCACT 1333
Qy 1278 TCTGGGAAGATTGCCAGCAGCAGAGGTGACGGTGTTCAGTACATTGGGAGCTGTGCC 1337
Db 1334 TCTGGGAAGATTGCCAGCAGCAGAGGTGACGGTGTTCAGTACATTGGGAGCTGTGCC 1333
Qy 1338 GATACCTTGTCAACGAGCCCGGAGCAAGGCAAGCTGGCCATAGGTCCGGCTGGCAG 1397
Db 1394 GATACCTTGTCAACGAGCCCGGAGCAAGGCAAGCTGGCCATAGGTCCGGCTGGCAG 1453
Qy 1398 TGGGAGCGGGGCTGGCCCAATACCTGGGAGCGTTTGTGGGGGCTTCGGGCGCCCTGC 1457
Db 1454 TGGGAGCGGGGCTGGCCCAATACCTGGGAGCGTTTGTGGGGGCTTCGGGCGCCCTGC 1513
Qy 1458 AGGTGCTGAGACATATGAGCTGACAGAGGGCAAGTGGCCACATCAATACACAGAC 1517
Db 1514 AGGTGCTGAGACATATGAGCTGACAGAGGGCAAGTGGCCACATCAATACACAGAC 1573
Qy 1518 ACGGGGGCGCTGGGGGCGTGTCTTCTGCTTACAGCATATCTTCCCTTCTCTTGA 1577
Db 1574 ACGGGGGCGCTGGGGGCGTGTCTTCTGCTTACAGCATATCTTCCCTTCTCTTGA 1633
Qy 1578 TTGGCTATGATGTCACACAGAGAGCCAAATTGGGACCCCGAGGGGCACTGTATGAGCA 1637
Db 1634 TTGGCTATGATGTCACACAGAGAGCCAAATTGGGACCCCGAGGGGCACTGTATGAGCA 1693
Qy 1638 CATCTCCAGGTAGCCAGGGGCTGTGTGTGCGCCCGGTAAAGCCAGCATCTCCCATTTCTGG 1697
Db 1694 CATCTCCAGGTAGCCAGGGGCTGTGTGTGCGCCCGGTAAAGCCAGCATCTCCCATTTCTGG 1753
Qy 1698 GGTATGCTGGCGGGCCAGAGCTGGGCCAGGGGAATGCTTAAGGATGTCTTCCGGGCTG 1757
Db 1754 GGTATGCTGGCGGGCCAGAGCTGGGCCAGGGGAATGCTTAAGGATGTCTTCCGGGCTG 1813
Qy 1758 GGGATGTTTTCTTCAACACTGGGGGACTGCTGTGTGCTGCGATGACCAAGTTTCTCCGCT 1817
Db 1814 GGGATGTTTTCTTCAACACTGGGGGACTGCTGTGTGCGATGACCAAGTTTCTCCGCT 1873
Qy 1818 TCCATGATGATGTCAGAGACACCTTCAAGTGGAAAGGGGAGAAATGTGGCCACACCGAGG 1877
Db 1874 TCCATGATGATGTCAGAGACACCTTCAAGTGGAAAGGGGAGAAATGTGGCCACACCGAGG 1933
Qy 1878 TGGCAGAGGCTTCTGAGGCGCTAGATTCTTCAAGAGGTGAAGCTATGAGTCACTG 1937
Db 1934 TGGCAGAGGCTTCTGAGGCGCTAGATTCTTCAAGAGGTGAAGCTATGAGTCACTG 1993
Qy 1938 TGGCAGAGGCTTCTGAGGCGCTAGATTCTTCAAGAGGTGAAGCTATGAGTCACTG 1997
Db 1994 TGGCAGAGGCTTCTGAGGCGCTAGATTCTTCAAGAGGTGAAGCTATGAGTCACTG 2053
Qy 1998 TGGACCTTATGAGGCTTCAACCAAGTGTGAGAGCTTGAACCTTATGCGGGGCGCC 2057
Db 2054 TGGACCTTATGAGGCTTCAACCAAGTGTGAGAGCTTGAACCTTATGCGGGGCGCC 2113
Qy 2058 GATTCCTCAGGCTCCAGAGTCTTGGCCACCAAGAGACCTTCAACAGCAGAAAGTTC 2117
Db 2114 GATTCCTCAGGCTCCAGAGTCTTGGCCACCAAGAGACCTTCAACAGCAGAAAGTTC 2173
Qy 2118 GATGGCAATAGAGGCTTCCAGCCAGACCTGTGACCTGACCTGATGAGCTTGAAGC 2177
Db 2174 GATGGCAATAGAGGCTTCCAGCCAGACCTGTGACCTGACCTGATGAGCTTGAAGC 2233
Qy 2178 AGGCTGATAGGCTTCACTGCGCTCACAACCTGCGGGTCAAGCGCCCTCTGGCAGGAA 2237
Db 2234 AGGCTGATAGGCTTCACTGCGCTCACAACCTGCGGGTCAAGCGCCCTCTGGCAGGAA 2293
Qy 2238 AACTTCGATCTGAGAACTTCCACACTGAGGCACTTGAAGAGAACTTGTGGGGTGG 2237
Db 2294 AACTTCGATCTGAGAACTTCCACACTGAGGCACTTGAAGAGAACTTGTGGGGTGG 2353
Qy 2298 GGGCGGTGACAGGTACTGGGCTGTGAGGAGTCTTTCTATACAGAACTGCGGTCACT 2357

Db 2354 GGGCGGTGACAGGTACTGGGCTGTGAGGAGTCTTTCTATACAGAACTGCGGTCACT 2413
Qy 2358 ATTTGTAATTAATGTGCTGAGCTGATCCAGCTGTCTGACCTA 2404
Db 2414 ATTTGTAATTAATGTGCTGAGCTGATCCAGCTGTCTGACCTA 2460

RESULT 13
ADA24640
ID ADA24640 standard; cDNA; 2574 BP.
XX
AC ADA24640;
XX
DT 20-NOV-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO703 cDNA.
XX
KW Human; secreted and transmembrane protein; PRO; gene; ss; tissue typing;
KW chromosome identification; vaccine; cancer; retinal disorder;
KW sports-related joint disorder; osteoarthritis; rheumatoid arthritis;
KW wound healing; obesity; diabetes; hearing loss;
KW cardiac insufficiency disorder; kidney disorder; nervous system disorder;
KW haemoglobin associated disorder.
XX
OS Homo sapiens.
XX
PN US2003050241-A1.
XX
PD 13-MAR-2003.
XX
PF 16-OCT-2001; 2001US-00978564.
XX
PR 17-OCT-1997; 97US-0062250P.
PR 03-NOV-1997; 97US-0064249P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066364P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077641P.
PR 11-MAR-1998; 98US-0077649P.
PR 12-MAR-1998; 98US-0077791P.
PR 13-MAR-1998; 98US-0078004P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078910P.
PR 20-MAR-1998; 98US-0078936P.
PR 20-MAR-1998; 98US-0078939P.
PR 25-MAR-1998; 98US-0079294P.
PR 26-MAR-1998; 98US-0079656P.
PR 27-MAR-1998; 98US-0079663P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079689P.
PR 27-MAR-1998; 98US-0079728P.
PR 27-MAR-1998; 98US-0079786P.
PR 30-MAR-1998; 98US-0079920P.
PR 30-MAR-1998; 98US-0079923P.
PR 31-MAR-1998; 98US-0080105P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080165P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080328P.
PR 01-APR-1998; 98US-0080332P.
PR 01-APR-1998; 98US-0080334P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 08-APR-1998; 98US-0081071P.
PR 09-APR-1998; 98US-0081195P.
PR 09-APR-1998; 98US-0081203P.
PR 09-APR-1998; 98US-0081229P.
PR 15-APR-1998; 98US-0081817P.
PR 15-APR-1998; 98US-0081819P.
PR 15-APR-1998; 98US-0081838P.
PR 15-APR-1998; 98US-0081952P.


```
PR 15-APR-1998; 98US-0081955P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082700P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 22-APR-1998; 98US-0082804P.
PR 23-APR-1998; 98US-0082796P.
PR 27-APR-1998; 98US-0083336P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083392P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083500P.
PR 29-APR-1998; 98US-0083545P.
PR 29-APR-1998; 98US-0083554P.
PR 29-APR-1998; 98US-0083558P.
PR 29-APR-1998; 98US-0083559P.
PR 30-APR-1998; 98US-0083742P.
PR 03-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 06-MAY-1998; 98US-0084419P.
PR 07-MAY-1998; 98US-0084598P.
PR 07-MAY-1998; 98US-0084600P.
PR 07-MAY-1998; 98US-0084627P.
PR 07-MAY-1998; 98US-0084637P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 13-MAY-1998; 98US-0085233P.
PR 13-MAY-1998; 98US-0085338P.
PR 13-MAY-1998; 98US-0085339P.
PR 15-MAY-1998; 98US-0085572P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085689P.
PR 15-MAY-1998; 98US-0085697P.
PR 15-MAY-1998; 98US-0085700P.
PR 15-MAY-1998; 98US-0085704P.
PR 18-MAY-1998; 98US-0086022P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086414P.
PR 22-MAY-1998; 98US-0086430P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087106P.
PR 28-MAY-1998; 98US-0087208P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 30-JUL-1998; 98US-0094651P.
PR 11-SEP-1998; 98US-0100038P.
PR 07-OCT-1998; 98WO-US021141.
PR 20-NOV-1998; 98US-0109304P.
PR 20-NOV-1998; 98WO-US024855.
PR 22-DEC-1998; 98US-0113296P.
PR 23-DEC-1998; 98US-0113621P.
PR 05-JAN-1999; 99WO-US000106.
PR 08-MAR-1999; 99WO-US005028.
PR 10-MAR-1999; 99WO-US005190.
PR 12-MAR-1999; 99US-0123957P.
PR 29-MAR-1999; 99US-0126773P.
PR 21-APR-1999; 99US-0130232P.
PR 28-APR-1999; 99US-0131022P.
PR 28-APR-1999; 99US-0131445P.
PR 14-MAY-1999; 99US-0134287P.
PR 14-MAY-1999; 99WO-US010733.
PR 02-JUN-1999; 99WO-US012252.
PR 16-JUN-1999; 99US-0139557P.
PR 23-JUN-1999; 99US-0141037P.
PR 07-JUL-1999; 99US-0142680P.

PR 26-JUL-1999; 99US-0145698P.
PR 28-JUL-1999; 99US-0146222P.
PR 29-OCT-1999; 99US-0162506P.
PR 30-NOV-1999; 99WO-US028313.
PR 02-DEC-1999; 99WO-US028551.
PR 02-DEC-1999; 99WO-US028565.
PR 16-DEC-1999; 99WO-US030095.
PR 30-DEC-1999; 99WO-US031243.
PR 30-DEC-1999; 99WO-US031274.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000277.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 24-FEB-2000; 2000WO-US005004.
PR 24-FEB-2000; 2000WO-US005841.
PR 10-MAR-2000; 2000WO-US006319.
PR 21-MAR-2000; 2000WO-US007532.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 24-AUG-2000; 2000WO-US023328.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001WO-US006520.
PR 22-MAR-2001; 2001WO-US009552.
PR 25-MAY-2001; 2001WO-US017092.
PR 01-JUN-2001; 2001WO-US017800.
PR 20-JUN-2001; 2001WO-US019692.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 30-JUL-2001; 2001US-00918585.

PR PA (GETH ) GENENTECH INC.
PI Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Baton DL;
PI Perrana N, Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME;
PI Goddard A, Godowski PJ, Grimaldi JC, Guirney AL, Hillan KJ,
PI Kijavitt IU, Kuo SS, Napier WA, Pan J, Paoni NF, Roy MA, Shelton DL;
PI Stewart TA, Tumas D, Williams PM, Wood WI;
XX WPI; 2003-521814/49.
XX P-PSDB; ADA24641.
XX DR
XX PT New isolated PRO polypeptides for example extracellular, secreted and
XX PT membrane bound proteins, useful for modulating the biological activities
XX PT of cells and for treating, for example diabetes, cancer, rheumatoid
XX PT arthritis, and hearing loss.
XX PT Claim 2; Fig 38; 461p; English.
XX XX
XX CC The invention describes an isolated secreted and transmembrane (PRO)
XX CC polypeptide (I). PRO337 polypeptide is useful for detecting PRO493
XX CC polypeptide in a sample, and vice versa. PRO725, PRO700 and PRO739 are
XX CC useful for detecting PRO1559 polypeptide in a sample, and PRO1559 is
XX CC useful for detecting PRO725, PRO700 and PRO739 in a sample. PRO493 is
XX CC useful for linking a bioactive molecule to a cell expressing a PRO337
XX CC polypeptide, and PRO337 is useful for linking a bioactive molecule to a
XX CC cell expressing a PRO493 polypeptide. PRO1559 is useful for linking a
XX CC bioactive molecule to a cell expressing a PRO735, PRO700 and PRO739
XX CC

Query Match 99.3%; Score 2387; DB 8; Length 2574;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 CTGCTCTCCGCGCCCTGTGAGTGTGGGGCCCTGGGTGGGATGGGCGTGTCCGAGCGA 77
Db 74 CTGCTCTCCGCGCCCTGTGAGTGTGGGGCCCTGGGTGGGATGGGCGTGTCCGAGCGA 133
QY 78 CGCGCGCTCCCTGGAAGAGAACTCTACGCTAGAACGAGCGGCCCTTAGGTTTCGGAAGG 137
```


QY 1158 AAGATGTGATCTACCTCGCCCTCCCACTTACACATGTCGGGTTCCCTGTCGGGATCG 1217
DB 1214 AAGATGTGATCTACCTCGCCCTCCCACTTACACATGTCGGGTTCCCTGTCGGGATCG 1273
QY 1218 TGGGCTGCATGGGCACTTGGGGCCCAAGTGTGCTAAATCCAAAGTTCTCGCTGGTCAAGT 1277
DB 1274 TGGGCTGCATGGGCACTTGGGGCCCAAGTGTGCTAAATCCAAAGTTCTCGCTGGTCAAGT 1333
QY 1278 TCTGGGAAGATTGGCAGCAGCAGGAGTGAAGGTTCTCCAGTACCTTGGGGAGCTGTGCC 1337
DB 1334 TCTGGGAAGATTGGCAGCAGCAGGAGTGAAGGTTCTCCAGTACCTTGGGGAGCTGTGCC 1393
QY 1338 GATACCTTGTCAACCAAGCCCGGAGCAGAAAGTGGCCATTAAGGTCGGCTGGCAG 1397
DB 1394 GATACCTTGTCAACCAAGCCCGGAGCAGAAAGTGGCCATTAAGGTCGGCTGGCAG 1453
QY 1398 TGGGCAAGGGGCTGGGCCCCAGATACCTGGGAGCCCTTTGTGGGGGCTTGGGGCCCTGCG 1457
DB 1454 TGGGCAAGGGGCTGGGCCCCAGATACCTGGGAGCCCTTTGTGGGGGCTTGGGGCCCTGCG 1513
QY 1458 AGGTCTGGAGACATATGACTGACAGAGGGCAAGTGGCCACATCAACTACACAGAGC 1517
DB 1514 AGGTCTGGAGACATATGACTGACAGAGGGCAAGTGGCCACATCAACTACACAGAGC 1573
QY 1518 AGCGGGGGCTGTGGGGGCTGTCTTCTGCTTTACAGCATATCTTCCCTTCTCTTGA 1577
DB 1574 AGCGGGGGCTGTGGGGGCTGTCTTCTGCTTTACAGCATATCTTCCCTTCTCTTGA 1633
QY 1578 TTGCTATGATGTCAACACAGAGAGCCAAATTCGGGAGCCCGGAGGGGCACTGTATGGCCA 1637
DB 1634 TTGCTATGATGTCAACACAGAGAGCCAAATTCGGGAGCCCGGAGGGGCACTGTATGGCCA 1693
QY 1638 CATCTCCAGGTGAGCAGAGGCTGTGTGTGTCGCCCGGTAAAGCCAGCATGCCATTCCTG 1697
DB 1694 CATCTCCAGGTGAGCAGAGGCTGTGTGTGTCGCCCGGTAAAGCCAGCATGCCATTCCTG 1753
QY 1698 GGTATGCTGGGGGCGGAGGCTGGGCGGCGGAGGAGATTTGCTAAAGATCTTTCGGGCTG 1757
DB 1754 GGTATGCTGGGGGCGGAGGCTGGGCGGCGGAGGAGATTTGCTAAAGATCTTTCGGGCTG 1813
QY 1758 GGGATGTTTTCTTCAACACTGGGGAAGCTGTGCTGTGCGATGACCAAGTTTTCTCCGCT 1817
DB 1814 GGGATGTTTTCTTCAACACTGGGGAAGCTGTGCTGTGCGATGACCAAGTTTTCTCCGCT 1873
QY 1818 TCCATGATCTGATCTGAGACACCTTCAGGTGGAAGGGGGAAGATGTGGCCAAACCGAGG 1877
DB 1874 TCCATGATCTGATCTGAGACACCTTCAGGTGGAAGGGGGAAGATGTGGCCAAACCGAGG 1933
QY 1878 TGGCAGAGGCTCTTCAGAGCCCTAGATTTCTTTCAGAGAGGTGAACGTATGAGTCACTG 1937
DB 1934 TGGCAGAGGCTCTTCAGAGCCCTAGATTTCTTTCAGAGAGGTGAACGTATGAGTCACTG 1993
QY 1938 TGGCAGGGCAAGAGGCAAGGCTGGAATGGCAAGCCCTAGTTCTGTGCTCCCCCAAGCTT 1997
DB 1994 TGGCAGGGCAAGAGGCAAGGCTGGAATGGCAAGCCCTAGTTCTGTGCTCCCCCAAGCTT 2053
QY 1998 TGGACCTTATGAGCTACACCCAGGCTGTGAGAACTTGCAACCTTATCCCGGAGCCG 2057
DB 2054 TGGACCTTATGAGCTACACCCAGGCTGTGAGAACTTGCAACCTTATCCCGGAGCCG 2113
QY 2058 GATTCCTCAGGCTCAGAGAGTCTTTGGCCACACAGAGACCTTCAACAGCAGAAAGTTC 2117
DB 2114 GATTCCTCAGGCTCAGAGAGTCTTTGGCCACACAGAGACCTTCAACAGCAGAAAGTTC 2173
QY 2118 GGAATGGCAAAATGAGGGCTTTCACCCAGACCTGTGACCCACTGATACCTTTTGAGACC 2177
DB 2174 GGAATGGCAAAATGAGGGCTTTCACCCAGACCTGTGACCCACTGATACCTTTTGAGACC 2233
QY 2178 AGGCTGTAGTGTCTTACCTGCGCCCTCACAACTGCGGCTACAGGCGCTCTCTGACAGAA 2237
DB 2234 AGGCTGTAGTGTCTTACCTGCGCCCTCACAACTGCGGCTACAGGCGCTCTCTGACAGAA 2293

QY 2238 ACCTTCGATCTGAGAACTTCCACACTGAGGACCTGAGAGAGAACTGTGGGGTGG 2297
DB 2294 ACCTTCGATCTGAGAACTTCCACACTGAGGACCTGAGAGAGAACTGTGGGGTGG 2353
QY 2298 GGGCCGTTGAGAGGTGATCTGGGCTGTGAGGATCTTTCTATACCAAGAACTGCGGTCACT 2357
DB 2354 GGGCCGTTGAGAGGTGATCTGGGCTGTGAGGATCTTTCTATACCAAGAACTGCGGTCACT 2413
QY 2358 ATTTGTAAATAATGAGCTGAGCTGATCCAGCTGTCTGACCTA 2404
DB 2414 ATTTGTAAATAATGAGCTGAGCTGATCCAGCTGTCTGACCTA 2460
RESULT 15
ADAI2301
ID ADAI2301 standard; cDNA; 2574 BP.
XX
AC ADAI2301;
XX
DT 06-NOV-2003 (first entry)
XX
DE Human cDNA encoding secreted/transmembrane polypeptide PRO703.
XX
KW ss; gene; inflammatory disease; organ failure; atherosclerosis;
KW cardiac injury; infertility; birth defect; premature aging; AIDS; cancer;
KW diabetic complication; tissue typing; human.
XX
OS Homo sapiens.
XX
FN US2003055216-A1.
XX
PD 20-MAR-2003.
XX
PF 17-OCT-2001; 2001US-00978824.
XX
PR 21-MAY-1996; 96US-0018049P.
PR 17-OCT-1997; 97US-0062250P.
PR 03-NOV-1997; 97US-0064249P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066364P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077641P.
PR 11-MAR-1998; 98US-0077649P.
PR 12-MAR-1998; 98US-0077791P.
PR 13-MAR-1998; 98US-0078004P.
PR 17-MAR-1998; 98US-00040220.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078910P.
PR 20-MAR-1998; 98US-0078936P.
PR 20-MAR-1998; 98US-0078939P.
PR 25-MAR-1998; 98US-0079294P.
PR 26-MAR-1998; 98US-0079656P.
PR 27-MAR-1998; 98US-0079663P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079689P.
PR 27-MAR-1998; 98US-0079728P.
PR 27-MAR-1998; 98US-0079786P.
PR 30-MAR-1998; 98US-0079920P.
PR 30-MAR-1998; 98US-0079923P.
PR 31-MAR-1998; 98US-0080105P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080165P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080328P.
PR 01-APR-1998; 98US-0080333P.
PR 01-APR-1998; 98US-0080334P.
PR 01-APR-1998; 98US-0080370P.
PR 08-APR-1998; 98US-0081070P.
PR 08-APR-1998; 98US-0081071P.
PR 09-APR-1998; 98US-0081195P.
PR 09-APR-1998; 98US-0081203P.
PR 09-APR-1998; 98US-0081229P.

```
PR 15-APR-1998; 98US-0081817P;
PR 15-APR-1998; 98US-0081819P;
PR 15-APR-1998; 98US-0081838P;
PR 15-APR-1998; 98US-0081952P;
PR 15-APR-1998; 98US-0081955P;
PR 21-APR-1998; 98US-0082568P;
PR 21-APR-1998; 98US-0082569P;
PR 22-APR-1998; 98US-0082700P;
PR 22-APR-1998; 98US-0082704P;
PR 22-APR-1998; 98US-0082797P;
PR 22-APR-1998; 98US-0082804P;
PR 23-APR-1998; 98US-0082796P;
PR 27-APR-1998; 98US-0083336P;
PR 28-APR-1998; 98US-0083322P;
PR 29-APR-1998; 98US-0083392P;
PR 29-APR-1998; 98US-0083495P;
PR 29-APR-1998; 98US-0083496P;
PR 29-APR-1998; 98US-0083499P;
PR 29-APR-1998; 98US-0083500P;
PR 29-APR-1998; 98US-0083545P;
PR 29-APR-1998; 98US-0083554P;
PR 29-APR-1998; 98US-0083558P;
PR 29-APR-1998; 98US-0083559P;
PR 29-APR-1998; 98US-0083742P;
PR 05-MAY-1998; 98US-0084366P;
PR 06-MAY-1998; 98US-0084414P;
PR 06-MAY-1998; 98US-0084441P;
PR 07-MAY-1998; 98US-0084598P;
PR 07-MAY-1998; 98US-0084600P;
PR 07-MAY-1998; 98US-0084627P;
PR 07-MAY-1998; 98US-0084637P;
PR 07-MAY-1998; 98US-0084639P;
PR 07-MAY-1998; 98US-0084640P;
PR 07-MAY-1998; 98US-0084643P;
PR 13-MAY-1998; 98US-0085323P;
PR 13-MAY-1998; 98US-0085338P;
PR 13-MAY-1998; 98US-0085339P;
PR 15-MAY-1998; 98US-0085573P;
PR 15-MAY-1998; 98US-0085579P;
PR 15-MAY-1998; 98US-0085580P;
PR 15-MAY-1998; 98US-0085582P;
PR 15-MAY-1998; 98US-0085689P;
PR 15-MAY-1998; 98US-0085697P;
PR 15-MAY-1998; 98US-0085700P;
PR 15-MAY-1998; 98US-0085704P;
PR 18-MAY-1998; 98US-0086023P;
PR 22-MAY-1998; 98US-0086392P;
PR 22-MAY-1998; 98US-0086414P;
PR 22-MAY-1998; 98US-0086430P;
PR 22-MAY-1998; 98US-0086486P;
PR 28-MAY-1998; 98US-0087098P;
PR 28-MAY-1998; 98US-0087106P;
PR 28-MAY-1998; 98US-0087208P;
PR 26-JUN-1998; 98US-0010541P;
PR 26-JUN-1998; 98US-0090863P;
PR 01-JUL-1998; 98US-0091010P;
PR 01-JUL-1998; 98US-0091359P;
PR 30-JUL-1998; 98US-0094651P;
PR 11-SEP-1998; 98US-0100038P;
PR 07-OCT-1998; 98US-0016697P;
PR 07-OCT-1998; 98WO-US02114P;
PR 06-NOV-1998; 98US-0018421P;
PR 06-NOV-1998; 98US-0018736P;
PR 20-NOV-1998; 98US-0109304P;
PR 20-NOV-1998; 98WO-US02485P;
PR 07-DEC-1998; 98US-0020205P;
PR 22-DEC-1998; 98US-0021851P;
PR 22-DEC-1998; 98US-0113296P;
PR 23-DEC-1998; 98US-0113621P;
PR 05-JAN-1999; 99WO-US00010P;
PR 08-MAR-1999; 99US-0025446P;
PR 08-MAR-1999; 99WO-US00502P;
PR 10-MAR-1999; 99US-0026566P;

PR 10-MAR-1999; 99WO-US00519P;
PR 12-MAR-1999; 99US-0026721P;
PR 12-MAR-1999; 99US-0123957P;
PR 29-MAR-1999; 99US-0126773P;
PR 12-APR-1999; 99US-0028429P;
PR 21-APR-1999; 99US-0130232P;
PR 26-APR-1999; 99US-0131022P;
PR 28-APR-1999; 99US-0131455P;
PR 14-MAY-1999; 99US-0031183P;
PR 14-MAY-1999; 99US-0134287P;
PR 14-MAY-1999; 99WO-US01073P;
PR 02-JUN-1999; 99WO-US01225P;
PR 15-JUN-1999; 99US-0139557P;
PR 23-JUN-1999; 99US-0141037P;
PR 07-JUL-1999; 99US-0142680P;
PR 26-JUL-1999; 99US-0145698P;
PR 28-JUL-1999; 99US-0146222P;
PR 25-AUG-1999; 99US-0038013P;
PR 25-AUG-1999; 99US-0038013P;
PR 25-AUG-1999; 99US-0038014P;
PR 29-OCT-1999; 99US-0162506P;
PR 30-NOV-1999; 99WO-US02831P;
PR 02-DEC-1999; 99WO-US02851P;
PR 02-DEC-1999; 99WO-US02855P;
PR 16-DEC-1999; 99WO-US03009P;
PR 30-DEC-1999; 99WO-US03124P;
PR 03-JAN-2000; 99WO-US00021P;
PR 06-JAN-2000; 99WO-US00027P;
PR 11-FEB-2000; 99WO-US00356P;
PR 18-FEB-2000; 99WO-US00434P;
PR 24-FEB-2000; 99WO-US00500P;
PR 02-MAR-2000; 99WO-US00584P;
PR 10-MAR-2000; 99WO-US00631P;
PR 21-MAR-2000; 99WO-US00753P;
PR 30-MAR-2000; 99WO-US00843P;
PR 17-MAY-2000; 99WO-US01370P;
PR 22-MAY-2000; 99WO-US01404P;
PR 30-MAY-2000; 99WO-US01494P;
PR 02-JUN-2000; 99WO-US01526P;
PR 28-JUN-2000; 99WO-US02071P;
PR 24-AUG-2000; 99WO-US02338P;
PR 08-NOV-2000; 99WO-US02328P;
PR 27-NOV-2000; 99WO-US02374P;
PR 01-DEC-2000; 99WO-US03267P;
PR 20-DEC-2000; 99WO-US03456P;
PR 20-DEC-2000; 99WO-US03456P;
PR 28-FEB-2001; 99WO-US00652P;
PR 22-MAR-2001; 99WO-US01674P;
PR 22-MAR-2001; 99WO-US01692P;
PR 10-MAY-2001; 99WO-US00505P;
PR 10-MAY-2001; 99WO-US00542P;
PR 10-MAY-2001; 99WO-US00542P;
PR 21-MAY-2001; 99WO-US01709P;
PR 01-JUN-2001; 99WO-US07203P;
PR 01-JUN-2001; 99WO-US01780P;
PR 05-JUN-2001; 99WO-US01750P;
PR 14-JUN-2001; 99WO-US008636P;
PR 19-JUN-2001; 99WO-US008634P;
PR 20-JUN-2001; 99WO-US01969P;
PR 29-JUN-2001; 99WO-US02106P;
PR 09-JUL-2001; 99WO-US02173P;
PR 30-JUL-2001; 99WO-US02185P;

XX (GETH ) GENENTECH INC.
XX
XX Ashkenazi AJ, Baker KP, Botstein D, Deansyere L, Baton DJ,
PI Ferrara N, Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME,
Query Match 99.3%; Score 2387; DB 8; Length 2574;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```


OY	18	CTGCTCTCCGCCCCTGATGAGTGGTGGGGGCGCTCGGTTGGGANTGGGCTGTCGACGGCA	77
Db	74	CTGCTCTCCGCCCCTGATGAGTGGTGGGGGCGCTCGGTTGGGANTGGGCTGTCGACGGCA	133
OY	78	CGCGGCTCTCTGGAGAGAAAGTCTCACTAGTAAAGACGAGCGGCTTAAGTTTTCGGAAG	137
Db	134	CGCGGCTCTCTCGGAAGAGAAAGTCTCACTAGTAAAGACGAGCGGCTTAAGTTTTCGGAAG	193
OY	138	GAGGATCAGGGANTGATTGGGAGCGGCTGGACACAGACGGTGGCGATAGAGGAAGCGGGCT	197
Db	194	GAGGATCAGGGANTGATTGGGAGCGGCTGGACACAGACGGTGGCGATAGAGGAAGCGGGCT	253
OY	198	CCATGAGCTGCGCTCTCTGCTGCTGGCCCCCTGCTGCTGTGTCTACCGCTGCTGTGAAGC	257
Db	254	CCATGAGCTGCGCTCTCTGCTGCTGGCCCCCTGCTGCTGTGTCTACCGCTGCTGTGAAGC	313
OY	258	TACACCTTCGCGCGGAGTGGGCTGGGCTTCGCGCGGACTTGGGCTTTGGGGTGGCGGCTC	317
Db	314	TACACCTTCGCGCGGAGTGGGCTGGGCTTCGCGCGGACTTGGGCTTTGGGGTGGCGGCTC	373
OY	318	TGTGCTGCAGAAAGGAGCTTTCGAGCTCGCGCCCTGAGCTGCGGACTGCGCTGACCTCGGAAG	377
Db	374	TGTGCTGCAGAAAGGAGCTTTCGAGCTCGCGCCCTGAGCTGCGGACTGCGCTGACCTCGGAAG	433
OY	378	GTCCCCGAGGGGGGCTGACGCTGAGCCTGGCGGCTCGCGGAACTGGGCCAGAGCGCGCG	437
Db	434	GTCCCCGAGGGGGGCTGACGCTGAGCCTGGCGGCTCGCGGAACTGGGCCAGAGCGCGCG	493
OY	438	CGCACACTTTTCTCATTTCACTGAGGCTCGCGGCGGCTTTAGCTACTCAGAGCGGAGCGCGAGA	497
Db	494	CGCACACTTTTCTCATTTCACTGAGGCTCGCGGCGGCTTTAGCTACTCAGAGCGGAGCGCGAGA	553
OY	498	GTAAAGAGGGCTGACAGCGGCTTCTCATGAGGCTTAGGGCTGGGAACTGGGGGACCGGACGGG	557
Db	554	GTAAAGAGGGCTGACAGCGGCTTCTCATGAGGCTTAGGGCTGGGAACTGGGGGACCGGACGGG	613
OY	558	GCACAGCTGCGGAGGAGGAGCGCTGGAGAGAAAGCGAGCGGAGCGGCTCGGGAGCGCGAGTG	617
Db	614	GCACAGCTGCGGAGGAGGAGCGCTGGAGAGAAAGCGAGCGGAGCGGCTCGGGAGCGCGAGTG	673
OY	618	CAGCGGCGGGAAGCGGCGCGGAGTTTTCGCGGAGGAGGAGCGGTGCGGCAAGAGTGGAGGAG	677
Db	674	CAGCGGCGGGAAGCGGCGCGGAGTTTTCGCGGAGGAGGAGCGGTGCGGCAAGAGTGGAGGAG	733
OY	678	CCGCGGCTCTCTGTCACTGTGAGCAACTGTGAGCGCTGCTCTCCCGCTGAGCCCAAGT	737
Db	734	CCGCGGCTCTCTGTCACTGTGAGCAACTGTGAGCGCTGCTCTCCCGCTGAGCCCAAGT	793
OY	738	TTCTGTGGTCTGATTCGGGCTGGGCGCAAGGCGGCGTGGGCACTGGCTTTGGGCCACCG	797
Db	794	TTCTGTGGTCTGATTCGGGCTGGGCGCAAGGCGGCGTGGGCACTGGCTTTGGGCCACCG	853
OY	798	CCCTGCGCGCGGGGCCCTCCCTGTGCACTGTCTTCGAGACTTCGAGCGCGCGCGCTGTGTGC	857
Db	854	CCCTGCGCGCGGGGCCCTCCCTGTGCACTGTCTTCGAGACTTCGAGCGCGCGCGCTGTGTGC	913
OY	858	TGGCGGCAAGATTTCGGAGTCCCTGTGAGCGGAGCTGGCGCGGCTGAGAGGCAATGGGGC	917
Db	914	TGGCGGCAAGATTTCGGAGTCCCTGTGAGCGGAGCTGGCGCGGCTGAGAGGCAATGGGGC	973
OY	918	TCACACTGTGGGCTGCAAGGCCAGGAATCCACCTTGTCTTGAATTAGCGATTTGTGTGCTG	977
Db	974	TCACACTGTGGGCTGCAAGGCCAGGAATCCACCTTGTCTTGAATTAGCGATTTGTGTGCTG	1033
OY	978	AAGTTCGAGTGAATGGATGGGGCAGTGCAGAGTAACTCTCTTCCCCCCAGAGATTA	103
Db	1034	AAGTTCGAGTGAATGGATGGGGCAGTGCAGAGTAACTCTCTTCCCCCCAGAGATTA	109
OY	1038	CAGACAGCTGCTGTACATCTTCACTCTGGGCAACAAGGCTCTCCCAAGGCTGCTCGGA	1097
Db	1094	CAGACAGCTGCTGTACATCTTCACTCTGGGCAACAAGGCTCTCCCAAGGCTGCTCGGA	1153

QY	1098	TCAGCTCACTCAAGAATCTCTGGAAATGCGCAGGGCTCTTAATCACTGTGTGTGTGTCCACAGG	1157
Db	1154	TCAGTCACTTCAAGATCTCTGGAAATGCGCAGGGCTCTTAATCACTGTGTGTGTGTCCACAGG	1213
QY	1158	AAGATGTATCTACCTCGCCCTCCCACTCTACCAATGTCCGGTTCCTCGTGGGACATCG	1217
Db	1214	AAGATGTATCTACCTCGCCCTCCCACTCTACCAATGTCCGGTTCCTCGTGGGACATCG	1273
QY	1218	TGGGCTGCAATGGGCATTTGGGGCCACAGTGTGTCTGGAATCCAAAGTTCTCGGCTGTCAGT	1277
Db	1274	TGGGCTGTATGGGCATTTGGGGCCACAGTGTGTCTGGAATCCAAAGTTCTCGGCTGTCAGT	1333
QY	1278	TCCTGGAAAGATTGCGACAGACACAGGGGTGACGGTGTTCACAGTACATTGGGAGCTGTGTC	1337
Db	1334	TCCTGGAAAGATTGCGACAGACACAGGGGTGACGGTGTTCACAGTACATTGGGAGCTGTGTC	1393
QY	1338	GATACCTTGTCAACCAAGCCCTCCGACCAAGGCAGAAAGCTGTGCATTAAGTCTCGGCTGGCAG	1397
Db	1394	GATACCTTGTCAACCAAGCCCTCCGACCAAGGCAGAAAGCTGTGCATTAAGTCTCGGCTGGCAG	1453
QY	1398	TGGGACAGGGGCTGGGCCACAGATACCTGTGGAGCGTTTTGTGTGGGCGCTTGCGGGCCCTGTC	1457
Db	1454	TGGGACAGGGGCTGGGCCACAGATACCTGTGGAGCGTTTTGTGTGGGCGCTTGCGGGCCCTGTC	1513
QY	1458	AGGTGTCTGAGACATATGGACTGACAGAGGGCAACGTGGCCACCATCACTACACAGAAC	1517
Db	1514	AGGTGTCTGAGACATATGGACTGACAGAGGGCAACGTGGCCACCATCACTACACAGAAC	1573
QY	1518	AGCGGGGGGCTGTGGGGCGTGTCTTCTGTGCTTTCAAGCATATTTCCCTCTCTCTTGA	1577
Db	1574	AGCGGGGGGCTGTGGGGCGTGTCTTCTGTGCTTTCAAGCATATTTCCCTCTCTCTTGA	1633
QY	1578	TTCCGCTATGATGTCAACCAAGGAGAGCCAAATCTGGAGACCCCAAGGGCACTGTATGSCCA	1637
Db	1634	TTCCGCTATGATGTCAACCAAGGAGAGCCAAATCTGGAGACCCCAAGGGCACTGTATGSCCA	1693
QY	1638	CATCTCCAGGTAGCCACAGGCTGTCTGTGTGSCCCCGTAAAGCCAGACAGTCCCAATTTCTGTG	1697
Db	1694	CATCTCCAGGTAGCCACAGGCTGTCTGTGTGSCCCCGTAAAGCCAGACAGTCCCAATTTCTGTG	1753
QY	1698	GCTATGCTGTGGGGGCGAGAGCTGTGSCCCAGGGGGAAGTTGCTAAAGATGTCTTCGGGCGTG	1757
Db	1754	GCTATGCTGTGGGGGCGAGAGCTGTGSCCCAGGGGGAAGTTGCTAAAGATGTCTTCGGGCGTG	1813
QY	1758	GCGATGTTTTCTTCAACAACCTGGGGACCTGTGTGTCTGTGCAATGACCAAGTTTTCCTCGCT	1817
Db	1814	GCGATGTTTTCTTCAACAACCTGGGGACCTGTGTGTCTGTGCAATGACCAAGTTTTCCTCGCT	1873
QY	1818	TCCATGATTCGATCTGAGAGACACTTTCAGGTGTGAAAGGGGAGAAATGTGGCCACACCGAAG	1877
Db	1874	TCCATGATTCGATCTGAGAGACACTTTCAGGTGTGAAAGGGGAGAAATGTGGCCACACCGAAG	1933
QY	1878	TGAGAGAGGTCTTGTGAAGGCCCTTAATTTCTTCAGAGAGTGAACGTGTATGAGATCACTG	1937
Db	1934	TGAGAGAGGTCTTGTGAAGGCCCTTAATTTCTTCAGAGAGTGAACGTGTATGAGATCACTG	1993
QY	1938	TGCCAGGGCATGAAGGCGAGGCTGGAAATGGCAGCCCTTAAGTTCTGTGCTCCCCCAGCGTT	1997
Db	1994	TGCCAGGGCATGAAGGCGAGGCTGGAAATGGCAGCCCTTAAGTTCTGTGCTCCCCCAGCGTT	2053
QY	1998	TGACACTTATCACTCTACACCCACAGCGTGTGGAACCTTGCCACCTTATGTGCCGGGCCC	2057
Db	2054	TGACACTTATCACTCTACACCCACAGCGTGTGGAACCTTGCCACCTTATGTGCCGGGCCC	2113
QY	2058	GATTCTCAGGCTCCAGAGGTCTTTTGGCCACCAACAGAGACTTTCAACACAGCAGAAAGTTC	2117
Db	2114	GATTCTCAGGCTCCAGAGGTCTTTTGGCCACCAACAGAGACTTTCAACACAGCAGAAAGTTC	2173
QY	2118	GGATGTGCAATGAAGGCTTGAACCCCAAGCAACCTGTGTGACCACTGTATGCTGTGAGC	2177
Db	2174	GGATGTGCAATGAAGGCTTGAACCCCAAGCAACCTGTGTGACCACTGTATGCTGTGAGC	2233
QY	2178	AGGCTGTAGTGCCTACCTGTGCCCTTCAACAATGCTCCGCTACAGGCGCCCTCTGTGGCAGGA	2237

```
Db      2234 AGGCTGTAGTGTGCTTACCTGCCCCCTCACAACTGCCGGTACAGCGCCCTCTGGCAGGAA 2293
QY      2238 ACCTTCGAATCTGAGAACTTCACACCTGAGGCACTGAGAGGAACTTGTGGGGTGG 2237
Db      2294 ACCTTCGAATCTGAGAACTTCACACCTGAGGCACTGAGAGGAACTTGTGGGGTGG 2353
QY      2298 GGGCCGTTGCAGGTGTACTGGGCTGTCAAGGATCTTTTCTATACCAAACTGCCGTCAC 2357
Db      2354 GGGCCGTTGCAGGTGTACTGGGCTGTCAAGGATCTTTTCTATACCAAACTGCCGTCAC 2413
QY      2358 ATTTGTATATAATGTGGCTGAGCTGATCCAGTGTCTCTGACCTA 2404
Db      2414 ATTTGTATATAATGTGGCTGAGCTGATCCAGTGTCTCTGACCTA 2460
```

Search completed: July 2, 2004, 23:28:03
Job time : 954 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2004, 17:32:25 ; Search time 21 Seconds

(without alignment)

Title: US-10-030-226-2

Perfect score: 3843

Sequence: 1 MGVCORTRAPFKESQLEA.....ALPLTTARYSALLAGLRI 730

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1257	32.7	689	1 JMW0107	very-long-chain ac
2	1056.5	27.5	646	1 AS5093	fatty acid transpo
3	1028	26.7	635	2 B87514	probable very-long
4	903	23.5	608	2 H83284	probable long-chain
5	858	22.3	655	2 T15879	hypothetical prote
6	841.5	21.9	597	2 D70609	probable fad6 pro
7	839.5	21.8	650	2 T21498	hypothetical prote
8	725.5	18.9	633	2 T43052	fatty acid transpo
9	654	17.0	623	2 S45899	probable membrane
10	405.5	10.6	522	2 H86484	probable crotonome
11	405.5	10.6	522	2 S40586	probable crotonome
12	393.5	10.2	522	2 S40586	probable carnitine
13	367	9.5	502	2 D70806	probable fad6 pr
14	353.5	9.2	517	2 A10510	probable crotonome
15	352	9.2	532	2 G70986	probable coa ligas
16	332	8.6	444	2 T50931	hypothetical prote
17	325	8.5	599	2 H72454	probable fatty-aci
18	313.5	8.2	632	2 B83163	probable AMP-bindi
19	311	8.1	584	2 C75364	probable long-chain
20	306	8.0	556	2 H81068	long-chain-fatty-a
21	302	7.9	507	2 G81798	long-chain-fatty-a
22	297.5	7.7	502	2 G70604	probable acid-CoA
23	295	7.6	564	2 C83327	probable AMP-bindi
24	293	7.5	508	2 H85271	probable long chain
25	291	7.6	541	2 T34850	probable acid-CoA
26	290.5	7.6	513	2 A69831	probable acid-CoA
27	286.5	7.5	515	2 F96011	probable long-chain
28	283	7.4	593	2 E69378	probable acid-CoA
29	278	7.2	1002	2 D87077	probable long-chain

30	276.5	7.2	562	2 AB0253	long-chain-fatty-a
31	273.5	7.1	561	2 B90743	long-chain-fatty-a
32	273.5	7.1	561	2 B85791	long-chain-fatty-a
33	273	7.1	524	2 H75530	probable acid-CoA
34	272.5	7.1	503	2 E70853	probable acid-CoA
35	269.5	7.0	561	1 S41589	long-chain-fatty-a
36	269.5	7.0	999	2 B70501	hypothetical prote
37	269	7.0	510	2 AB2782	long-chain fatty a
38	269	7.0	510	2 C97561	hypothetical prote
39	269	7.0	514	2 T46131	4-coumarate-CoA 11
40	268	7.0	512	2 T41164	coenzyme a synthet
41	268	7.0	566	2 H82132	long-chain-fatty-a
42	267.5	7.0	639	2 AB3456	long-chain-fatty-a
43	265	6.9	562	2 B83233	long-chain-fatty-a
44	265.5	6.9	561	2 AC0725	long-chain-fatty-a
45	264	6.9	569	2 D82824	regulator of patho

ALIGNMENTS

RESULT 1

very-long-chain acyl-CoA synthetase related protein - mouse
N:Alternate names: VLACSR
C:Species: Mus musculus (house mouse)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 08-Sep-2000
C/Accession: JMW0107
R:Berger, U.; Truppe, C.; Neumann, H.; Fors-Petter, S.
Biochem. Biophys. Res. Commun. 247, 255-260, 1998
A>Title: A novel relative of the very-long-chain acyl-CoA synthetase and fatty acid tran
A/Reference number: JMW0107, PMID:98308102; PMID:9642112
A/Accession: JMW0107
A/Molecule type: mRNA
A/Residues: 1689 <BBR>
A/Cross-references: GB:AJ223959
A/Experimental source: liver
C/Comment: This protein likely functions as a plasma membrane transporter of long chain
cids.
C:Superfamily: Mycobacterium tuberculosis probable fad6 protein; acetate-CoA ligase hom
F;169-647/Domain: acetate-CoA ligase homology <AC1>

Query Match	32.7%	Score 1257;	DB 1;	Length 689;
Best Local Similarity	39.5%;	Pred. No. 3.8e-79;		
Matches 280;	Conservative 93;	Mismatches 257;	Indels 78;	Gaps 8;
QY	46	GSMAATLPLLL	-----LPLLLKIHMPQLRMLPADLAPVRA	86
DB	37	GDPTCLVLLGLLGRPWISSMMPHWSLVGAALTLFLPLPPLPPGLRLHNDVATFPKX	96	
QY	87	LCKKRALRATLAAAADPEGGCSLAWRLAELAQORAAHTFLIHGRRRFSYSAERE	146	
DB	97	LFYGLKPRRR-----LNKHPPE-----TFV-----DALERQ	122	
QY	147	SNMAAATFLALGMDMPDG-----GDSGSGAGEGERRAPAGADAAAGSAGEPAGDGA	202	
DB	123	-----ALAW-----PRVALVCTGSSSSSTTNSQLDARSQQAAMVTLAKKDKDAVTON	169	
QY	203	RGGAAPLSPGATVALLPAGPEFLMLFWGLAKAGIRTAFTVPTALRGLHLCLHSCGA	262	
DB	170	TRMAALIVPSTITALS-----VFLGLAKGCPVAMINPHSRGMFLHSVSSGA	221	
QY	263	RALVLAPEFLSELPDLPALRAMGLHMAAGPETHAGISDLAEVSAEVDGVPYGLSS	322	
DB	222	SVIIVDPDQENIEVLPKILAEINHCFLYGHSSPTPGVEALGASLDAAPSDVPVSLGA	281	
QY	323	PGSITDCLVYIFTSGTGLPKARISHLKILOCGGFYOLCGVHQEDVITLALPYLMSGS	382	
DB	282	TIRKSPATIFITSGTGLPKRALISHERVIVSNVLSGCGADVVVDVPLVITIDL	341	
QY	383	LIGIVGGMIGATVIVKSRFACQFEDCQQRVTVFYIYGLCRYLVNQPSSKAEGRHK	442	
DB	342	VLFGLGLQVATGATCVLAPKFSASRFWAECROHGVTVILVGEITRLYLCNVPPQEDKIH	401	

Cy	572	NTGGLVLCDDQGGFLRPHDRGTDFFRWKGENVAATTEVAEYFALDPLFOENNVVGATYPGHE	631
Db	489	LSGDVLIMDELGVWYFFRDRSGDTFRMRGENVSFTTEBAVLSRLGGOTDAVAVGAAPGVE	548
Cy	632	GRAGMAALVLRPPHA-LIDLMQLYTHVSENLPPIARPRFRLQSLAATTEFFRKOKXRMAN	690
Db	549	GKAGMAAIA--DPHSQIDPNSMYCELOKVLAASARPIFKLLFPVDITGTGIKTORLOR	606
Cy	691	EGFDPTLSDPLYLDQAVGAYPLLTARYSALLAGNLRI	730
Db	607	EGFDPTRTSLRLEFFLDLKQGRVYPLDERVHARRCAGDFSL	646
 RESULT 3 B87534			
fatty acid transport protein, probable CC2299 [imported] - Caulobacter crescentus			
C:Species: Caulobacter crescentus			
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #ext_change 10-May-2001			
C:Accession: B87534			
C:Nierman, W.C.; Felblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heideberg, J.			
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.U.; Duxkin, A.S.; Gilm, M.L.; Hart, D.H.; Kolon			
n, U.; Ermolaeva, M.; White, O.; Saliberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.			
proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001			
A>Title: Complete Sequence of Caulobacter crescentus.			
A:Reference number: A87249; MUID:21173698; PMID:11259647			
A:Accession: B87534			
A>Status: preliminary			
A:Molecule type: DNA			
A:Residues: 1-635 <STO>			
A:Cross-references: GB:AE005673; NID:g13423818; PIDN:AAK24270.1; GSPDB:GN00148			
C:Genetics:			
A:Gene: CC2299			
C:Superfamily: Mycobacterium tuberculosis probable fadD6 protein; acetate-CoA ligase hom			
 Query Match 26.7%; Score 1028; DB 2; Length 635; Best Local Similarity 41.4%; Pred. No. 2,4e-53; Matches 23;/ Conservative 72; Mismatches 239; Indels 24; Gaps 9;			
Cy	162	WGPDGDSGSGAGGEGRARPGAGNDAAAGCAFGAGDGAARGGAAPLSPGATVALLL	221
Db	162	WGPDGDSGSGAGGEGRARPGAGNDAAAGCAFGAGDGAARGGAAPLSPGATVALLL	221

[illegible]

Db 547 LVTEGD--FDLEFAKYYDEALPSYARPLFVRLQKAIEFTTGFKYRKIDLVNNEGDSKST 604
QY 699 SPPLVYLDQAVGAYLPLTTARYSALLAGNLR 730
Db 605 KDPLFRDPKAG-YVKITKTICAKIEGGGFR 635

RESULT 4

probable very-long-chain acyl-CoA synthetase PA2893 [imported] - Pseudomonas aeruginosa
C/Species: Pseudomonas aeruginosa
C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C/Accession: H83284
R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warener, P.; Hickey, M.J.; Berman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Laidig, K.; Lim, J.; Lory, S.; Olson, M.V.
Nucleotide: 406, 959-964, 2000
A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A/Reference number: A82950; MUID:20437337; PMID:10984043
A/Accession: H83284
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-608 <STO>
A/Cross-references: GB:AB004715; GB:AB004091; NID:93948977; PIDN:AAG06281.1; GSPDB:GN001
A/Experimental source: strain PA01
C/Genetics:
A/Gene: PA2893
C/Superfamily: Mycobacterium tuberculosis probable fad6 protein; acetate-CoA ligase hom

Query Match 23.5%; Score 903; DB 2; Length 608;
Best Local Similarity 35.3%; Pred. No. 1e-54;
Matches 223; Conservative 80; Mismatches 243; Indels 86; Gaps 10;

QY 109 EGGCSLARLAEIAQRAHFTFLHGSRRFSSEARSRNARAPFLALGMDWGDPDGD 168
Db 39 EKSLSLGALRLARRALYDNDNAVLDGQRISYALFNGMANRLARF----- 84
QY 169 SGESEAGGERAPAGDAAAGSGAEFAGDGAARGGAAAPLSPGATVALLPAGDEFL 228
Db 85 -----KAEVGH-----GSVAVALLENREVEL 106
QY 229 WLMFGLAKGLRTAFTVPLRRGPLHCLRSOGARALVLAPELSELPDLPALRAGLH 288
Db 107 AILALALALGAILGALVNTTQKGLVHSLNIVKGFHVGELREVEEYHVLGNAGH 166
QY 289 LM-----AAGPETHPAGISDL-----AEVSAEVDGVPVGYLSPOSITDCLYIFT 335
Db 167 CYWVDDGDTLGDPPGSPFGWMENLMLRAGQTSNLEDTGRV-----RLKSCFYIYT 218
QY 336 SGTGLPRAARISHKILQCCGFF--QLCGVHOEDVITYLALPLYHMGSLIGIVCGKIG 393
Db 219 SGTGLPFAASIMSHKMTKAYGFGHSGILGARDVLTLLPCYHNNAVVTCMSAALAGG 278
QY 394 ATVVLKSFSAQGFEDCCQHRVTVFQYIGELCRVLYNQPPSKARSGHKVALVAGSGLR 453
Db 279 AAMALRRFSAISGFHKVDQHYRATCFYIGELCRVLYNQPPCAERKSLTCTMGNGIRP 338
QY 454 DTWERFARFEPLOVLEYGLTEGNVATINTGORGAVRASMLYKHIFPPLIRYDVT 513
Db 339 SIMAEFKRFEIORTITEFVASEGNIQFTNFFNDTVGFSPATY-----AIRYDLEN 392
QY 514 GEPTRDPQGMATSPGEGLLVAPVSOQSPFLYAGGPELAGKLLKDVPRPDVFNT 573
Db 393 DRPRDPAKFMERKGVKSGVGLLSEISKMPFDGT-DPAKSEAVILLRDFKGDVAFNT 451
QY 574 GDLLVCDQGF--LRFHRTGDTFRMKGENVATTEVAEVEALDFLOEVNVTGTVGHE 631
Db 452 GDLN--RDIGFKHQFVRLGDTFRMKGENVSTEVEALGAFDGVEDAVVYGVLEIPGN 509
QY 632 GRAGAAALVLRPPLALDMLQYTVHSENLPRYARPRFLRLOESLATTETPRQGVRRANE 691
Db 510 GRGMAALRLADGVLELSDALAAHLDRELPAVATPVFLRLREVEETGTFRYKKTDLKRD 569

QY 692 GFDPSTSDPLVYLDQAVGAYLPLTTARYSAL-723
Db 570 AYDPARVSDKLFVRLPESAGYQPLDAELXAL 601

RESULT 5

hypothetical protein D1009.1 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 02-Sep-2000
C/Accession: T15879
R/Favell, T.
submitted to the EMBL Data Library, November 1995
A/Description: The sequence of C. elegans cosmid D1009.
A/Reference number: Z18422
A/Accession: T15879
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-655 <FAV>
A/Cross-references: EMBL:U40938; NID:g1072169; PID:g1072173; PIDN:AAA81698.1; CESP:D1009
C/Genetics:
A/Gene: CESP:D1009.1
A/Intons: 2/2; 60/2; 89/3; 125/3; 192/1; 344/2; 399/1; 547/3; 569/3; 602/1
C/Superfamily: Mycobacterium tuberculosis probable fad6 protein; acetate-CoA ligase hom
F/128-611/Domain: acetate-CoA ligase homology <ACT>

Query Match 22.3%; Score 858; DB 2; Length 655;
Best Local Similarity 37.9%; Pred. No. 1.5e-51;
Matches 197; Conservative 95; Mismatches 200; Indels 28; Gaps 12;

QY 214 GATVALLPAGPEFLWLMFGLAKGLRTAFTVPLRRGPLHCLRSOGARALV-----LAP 269
Db 132 GDVVALFMENSIDFPALWGLSKIGVSAFINSNMLEPLASHINVSCKKSCITNINLP 191
QY 270 EFLSELPDLPALRAGLHMAAG--PETHPAGISDLAEVSAEVDGVPVGYLSPOSIT 326
Db 192 MFRAERKMLISDE--HVFVLAGTVDRHNSLQDD--HFSDEDPVIGL--NF 242
QY 327 TDTCLYFTSGTGLPRAARISHKILQ--CQGYVOCGVHOEDVITYLALPLYHMGSLIG 385
Db 243 RSLVCLYITSGTGNKRAVITKHFRTAMAGKAFGINKSDVYITPMHSAAGING 302
QY 386 IYVCMGIGATVLKSPFSAQGFEDCCQHRVTVFQYIGELCRVLYNQPPSKARSGHKVL 445
Db 303 IGSILIAFGSTAVVRRKFSANFWDCKVNVATAYIYGEICRYLLAANCPPEKQHVR 362
QY 446 AVSSGLRPTWERFARFEPLOVLEYGLTEGNVATINTGORGAVRASMLYKHIF--F 502
Db 363 MGAAGLRGQIWEKFEVRFKIGELYSGTEGNSIVNVDNHVAGGEMP-ITPHIGSLY 421
QY 503 PFLIRYDVTYTGEPFIRDPQGMATSPGEGLLVAPVSOQSP--PFLYAGGPELAGKLL 559
Db 422 PVRLITVDATGELERDKNGLCVCPVGETGEVGYIKDILKEGYVSGDTAK-KI 480
QY 560 LKDVPRPDVFNTGGLVLCDDQGPLRFRDRTGDTFRMKGENVATTEVAEVEALDFLOE 619
Db 481 YRVFNGHGXVFRASGGLHMDLGYLYFVDRCGDFFRMGENVSTTEVGLIQPVNDVAD 540
QY 620 VNVYGVTPGHEGRAGMALVLRPPLALD--LMQLYTHSENLPRYARPRFLRLOESLAT 677
Db 541 ATYGVTVKMGGRAGMAIVKDGVDKFIADITSRLTENLASAIVFIRLCEVDR 600
QY 678 TETFKQCKVRMANEGFD-BSTLSDPLVYLDQAVGAYLPLT 716
Db 601 TGFFLKKTDLQKGYDLVACKDPIYVWSAEKSYKPLT 640

RESULT 6

D70609
probable fad6 protein - Mycobacterium tuberculosis (strain H37RV)
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 18-Aug-2000
C/Accession: D70609

Query Match 21.9%; Score 841.5; DB 2; Length 597;
C/Superfamily: Mycobacterium tuberculosis probable fad6 protein; acetate-CoA ligase hom
F_93-555/Domain: acetate-CoA ligase homology <AC>

RESULT 7
T21498

A:Map position: 4
A:Introns: 55/2; 87/2; 120/3; 187/1; 241/1; 356/1; 355/1; 600/1
C:Superfamily: Mycobacterium tuberculosis probable fadD protein; acetate-CoA ligase homology
F123-609/Domain: acetate-CoA ligase homology <ACU>

QY 669 LRLESLATTETFKQKRYRANEGFDEPDSSTDELVIYLDAQVAYLP 714
::: : | ||| : : : : :
Db 590 MRICQDVYEKTGTFKLVAKTINLQRGLMDAP-SDSIYYINSNRNRPVP 634

A;Accession: T43052
A;Status: preliminary; translated from GB/EMBL/DBJ

C:Superfamily: 4-coumarate-CoA ligase; acetate-CoA ligase homology

Query Match 10.6%; Score 405.5; DB 2; Length 522;
Best Local Similarity 25.8%; Pred. No. 2,2e-20;
Matches 154; Conservative 78; Mismatches 235; Indels 131; Gaps 19;

```

QY 110 GGCSLAWRLAELAQCPAAHTFLIHGS-----RRFYSSEARERESNRPAAPFLRALGMDWGP 164
DB 10 GGCHLRQWMDLADYGHKTLALICSSGGVNNRYSTLEINQINNTANLFF-TLG----- 63
QY 165 DGDSDSGSAGEGERAAPGADMAAGSGAEFAGGGAAGGAAPLSPGATVALLPAG 224
DB 64 -----IRKGDVYALHLNDNC 77
QY 225 PEFLMFMGLAKGLATAVPTALRRGPLLHCLRSCGAPAVLA---PEFLSELEPDL 280
DB 78 PEFLFCMFLAKGLATVMPINARLLRESAMILQNSQACLVTSAQFYMYQIQOEDAT 137
QY 281 ALRAMGLHMAAGPGTHPAGISDLAEVSAEVDGPVGY-----LSSPOSTD 328
DB 138 QLR-----HI-----CLTDVALPADDVSSFTQKNOQPATLCYAPPLSTDD 179
QY 329 TCLYITSGTTGLPKAARISHLKILQCOGFYQL--CGVQEDVYIALPLYMSSGLIGI 386
DB 180 TAEILFTSGTSTRPKGVVITHYN-LRFAGYSAQCALRDVDVYLTVPAPAHIDCOCTAA 238
QY 387 VCGMGIGATVVLKSKFSAGQFWEQCOQHVTVFOYIGELCRVLYVNOPESKARGHKVR-L 445
DB 239 MAFSAGATFVLVEKFSAPAFMGQVOKYRATITTECPMMIRLTMQPSANDRHRLEEV 298
QY 446 AVSGSLRPPTWERFVRFPGLQVLETYGLTEGNVATINTYQGR-----GAVGRASWLY 498
DB 299 MYFLNLSQEKDTFCERFG-VRLTISYGMTETIVGII--GDRPGDKRRMPSIGRAGFCY 354
QY 499 KHIFPSSLIRYVTTGEPTRDPQGHCMATSPGEPG-LTVAIPSOQSPFLGAGPELAOG 557
DB 355 -----DAEIRL--DNRRLPAGEIGEICIKGYPKGKTIPEYFLNKR-ATA 396
QY 558 KLLKQVFRPDVFNFTGDLVLCDDQGFLEFHDRTGDTFRMKENAVTEVAEVEALDFL 617
DB 397 KYLE-----ADGMHTGDTGYRDEEGFFYFIDRCNMIRKGENVSCVLENIITATHPKI 451
QY 618 QEVNYYGVTVPGHEGAGMAALVLRPPHALDMQLYTHVSENLPYARPRFLRLOESL 675
DB 452 QDIVVVGIR-DSIRDEAIKAFVNLNEGFTLSEEFRCQNMAKFKVSYLEIRKDL 508

RESULT 11
H85484
probable crotonobetaine/carnitine-CoA ligase caic [imported] - Escherichia coli (strain
C/Species: Escherichia coli
C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C/Accession: H85484
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grochbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamocitis, K.; Apodaca,
Nature 409, 529-533, 2001
A/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A/Reference number: A85480; MUID:21074935; PMID:11206551
A/Accession: H85484
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-522 <STO>
A/Cross-references: GB:A8005174; NID:g12512724; PIDN:AAG54340.1; GSPDB:GN00145; UMGF:200
A/Experimental source: strain O157:H7, substrain EDJ933
C/Genetics:
A:Gene: caic
C:Superfamily: 4-coumarate-CoA ligase; acetate-CoA ligase homology

Query Match 10.6%; Score 405.5; DB 2; Length 522;
Best Local Similarity 25.8%; Pred. No. 2,2e-20;
Matches 154; Conservative 78; Mismatches 235; Indels 131; Gaps 19;
QY 110 GGCSLAWRLAELAQCPAAHTFLIHGS-----RRFYSSEARERESNRPAAPFLRALGMDWGP 164

```

```

DB 10 GGCHLRQWMDLADYGHKTLALICSSGGVNNRYSTLEINQINNTANLFF-TLG----- 63
QY 165 DGDSDSGSAGEGERAAPGADMAAGSGAEFAGGGAAGGAAPLSPGATVALLPAG 224
DB 64 -----IRKGDVYALHLNDNC 77
QY 225 PEFLMFMGLAKGLATAVPTALRRGPLLHCLRSCGAPAVLA---PEFLSELEPDL 280
DB 78 PEFLFCMFLAKGLATVMPINARLLRESAMILQNSQACLVTSAQFYMYQIQOEDAT 137
QY 281 ALRAMGLHMAAGPGTHPAGISDLAEVSAEVDGPVGY-----LSSPOSTD 328
DB 138 QLR-----HI-----CLTDVALPADDVSSFTQKNOQPATLCYAPPLSTDD 179
QY 329 TCLYITSGTTGLPKAARISHLKILQCOGFYQL--CGVQEDVYIALPLYMSSGLIGI 386
DB 180 TAEILFTSGTSTRPKGVVITHYN-LRFAGYSAQCALRDVDVYLTVPAPAHIDCOCTAA 238
QY 387 VCGMGIGATVVLKSKFSAGQFWEQCOQHVTVFOYIGELCRVLYVNOPESKARGHKVR-L 445
DB 239 MAFSAGATFVLVEKFSAPAFMGQVOKYRATITTECPMMIRLTMQPSANDRHRLEEV 298
QY 446 AVSGSLRPPTWERFVRFPGLQVLETYGLTEGNVATINTYQGR-----GAVGRASWLY 498
DB 299 MYFLNLSQEKDTFCERFG-VRLTISYGMTETIVGII--GDRPGDKRRMPSIGRAGFCY 354
QY 499 KHIFPSSLIRYVTTGEPTRDPQGHCMATSPGEPG-LTVAIPSOQSPFLGAGPELAOG 557
DB 355 -----DAEIRL--DNRRLPAGEIGEICIKGYPKGKTIPEYFLNKR-ATA 396
QY 558 KLLKQVFRPDVFNFTGDLVLCDDQGFLEFHDRTGDTFRMKENAVTEVAEVEALDFL 617
DB 397 KYLE-----ADGMHTGDTGYRDEEGFFYFIDRCNMIRKGENVSCVLENIITATHPKI 451
QY 618 QEVNYYGVTVPGHEGAGMAALVLRPPHALDMQLYTHVSENLPYARPRFLRLOESL 675
DB 452 QDIVVVGIR-DSIRDEAIKAFVNLNEGFTLSEEFRCQNMAKFKVSYLEIRKDL 508

RESULT 12
S40558
probable carnitine-CoA ligase (EC 6.2.1.-) - Escherichia coli (strain K-12)
C/Species: Escherichia coli
C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 01-Mar-2002
C/Accession: E64724; S40558; I41013
R:Blaetner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
Science 277, 1453-1462, 1997
A/Title: The complete genome sequence of Escherichia coli K-12.
A/Reference number: A64720; MUID:97426617; PMID:9278503
A/Accession: E64724
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-522 <BLAT>
A/Cross-references: GB:A8000114; GB:U00096; NID:g1786217; PIDN:AACT3148.1; PID:g1786221;
A/Experimental source: strain K-12, substrain MG1655
R:Ylstra, T.; Mort, H.; Nagai, H.; Nagata, T.; Ishihama, A.; Fujita, N.; Isono, K.; Mizobu
submitted to the EMBL Data Library, December 1992
A/Description: Systematic sequencing of the Escherichia coli genome: analysis of the 0-2
A/Reference number: S40531
A/Accession: S40558
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-522 <YLR>
A/Cross-references: EMBL:D10483; NID:g216434; PIDN:BA01313.1; PID:g216462
R:Elchler, K.; Bourgis, F.; Buchet, A.; Kleber, H.P.; Mandrand-Berthelot, M.A.
Mol. Microbiol. 13, 775-786, 1994
A/Title: Molecular characterization of the cai operon necessary for carnitine metabolism
A/Reference number: I41010; MUID:95115548; PMID:7815937
A/Accession: I41013
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA

```


A/Cross-references: GB:AL513382; PIDN:CAD01225.1; PID:G16501354; GSPDB:GN00176
 C/Genetics:
 A/Gene: SRY0081
 C/Superfamily: 4-coumarate-CoA ligase; acetate-CoA ligase homology

Query Match 9.2%; Score 353.5; DB 2; Length 517;
 Best Local Similarity 24.4%; Pred. No. 8.7e-17;
 Matches 144; Conservative 84; Mismatches 248; Indels 113; Gaps 18;

QY 110 GGCSLAWRLAEALQOAAATPLIHGS-----RRFSSEARESENAPAPALFALGMDWP 164
 DB 5 GGNLRQWMDDLAAGWGDGKTLALIFSCGIVNQFYSALBENIRNTALVF----- 55
 QY 165 DGDSDGSGSAGEERAPGADAAAGSGAEFAGDGAARGGGAAPSPGATVALLPAG 224
 DB 56 -----YLG-----IRKGRVALHLDNC 72
 QY 225 PEFMLWPGAKAGRTAFVPTALRRGPLLHCRSGARALYLAPEFL-----ESLEP 277
 DB 73 PEFITCFWFLAKIGALMPFIPNRLGEBASWILQMSQVSLVTSAQFPMREIRQDNT 132
 QY 278 DLPALRAMGLHMAAGPSTHPAGISDLAEVSAEVDGVPVGLSSPGSTDTCLYIFTSG 337
 DB 133 PLNHICLIGEQLPADDGVSH---FSQLQKQSATL-----CYTPALSTDDTAETILFTSG 183
 QY 338 TTGLPRAARISHLKITLQCGEYQL---CGVHEDVLYLALPLYHMSGSLGIVGCMGIGAT 395
 DB 184 TTSRPGVAVITHN-LRFAGYISAWOIALRDBDVTMTVPAPAFHICCOCTAAMPASAGST 242
 DB 396 VVLKSKFSAQGFWECCQHRVTVFQYIGELCRYLVNQPESKAERGHKVR-LAVSGLRPD 454
 QY 243 FVLEKXSAFAPMDQYRKYQATVTECIPWIRTLWQATPDROHIREVWFYNTLSAQ 302
 QY 455 THERFPRRGPLOVLETYGLTEGNATINYTGQ-----GAVGRASWLYKHIFPFLI 507
 DB 303 EKDAFEREPG-VALLSYVMTETTVGIT---GDRGDKRMPDSIRVGSTF----- 350
 QY 508 RYDVTGEPFRPOGHCMATSPGEPG-LTAPVSQSPFLGYAGPELAOGKLTMDVERP 566
 DB 351 -----AATRDQNRPLPA--GEIGELCIKGIPEKTIIFXEYMPPE-ATRAALP----- 396
 QY 567 GGVFENITGDLVCDGFLRFHRTGDFRMKGENVATTEVAEVALDFLOEVNVYGT 626
 DB 397 -EGWLTHTGDSGYODEDGYFYFVDRRCNMKRGSEVNSCYELNIIISAPKIODIYVVGIR 455
 QY 627 VPGHEBRAGMALVLEPPHALDLMOLYTVSENLPYPARPRFLQESL 675
 DB 456 -DAIRDEAIKAFIVLNEGELTSEAEFFSCENNAKFKVPSPMEIRTDL 503

RESULT 15

G70986
 Probable CoA ligase - Mycobacterium tuberculosis (strain H37RV)

C/Species: Mycobacterium tuberculosis
 C/Date: 17-Jul-1998 #sequence: revision 17-Jul-1998 #text_change 18-Aug-2000

C/Accession: G70986

R/Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998

A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A/Reference number: A70500; PMID:9829587; PMID:9634230

A/Accession: G70986

A/Status: Preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-532 <COL>

A/Cross-references: GB:Z95890; GB:AL123456; NID:G3242245; PIDN:CAM09316.1; PID:G2131015

A/Experimental source: strain H37RV

C/Genetics:

A/Gene: fadD1

C/Superfamily: 4-coumarate-CoA ligase; acetate-CoA ligase homology

F:48-499/Domain: acetate-CoA ligase homology <ACU>

Query Match 9.2%; Score 352; DB 2; Length 532;
 Best Local Similarity 27.2%; Pred. No. 1.1e-16;
 Matches 146; Conservative 76; Mismatches 244; Indels 70; Gaps 21;

QY 199 DGAARGGG---AAAPSPGATVALLPAPPEFLMIFGAKAGLTAFTAPFAL-----RRGP 252
 DB 35 ESAARPAALITLADPPRP-THIGSLIGNTPML---AQAAAGL-CGYLCCGLANTRRGD 89
 QY 253 LHC-LRSCGARALVLAPEFLSLEP-DLPALRANGLH--WA---AGGTTPAGISDL 305
 DB 90 ALADVRADDCQIVVDADHRAALDGLDLAGARILDTSTPRYAEVAGGAFV----- 143
 QY 306 AEVSAEVDGVPVGLSSPGSTDTCLYIFTSGLPKARISHL-KILQCGFYQLCGV 364
 DB 144 ---YREVD-----TMDPFMMITSTGSGNPKAVPVSHLMATAGSLIERPGL 188
 QY 365 HOEDVLYLALPLYHMSGSLGIVGCMGIGATVLLSKFSAGGFWECCQHRVTVFQYIG 424
 DB 189 TEGDTCYVSMPLFHSNAVVAWMAFVAGSAAIA-PATFSATGFLDVRRYHATYMYVGK 247
 QY 425 LCRYLVNQPESKAERGHKRLAVGSGLRPDTERFVRRFGPLOVLETYGLTEGNATINY 484
 DB 248 PLAVIATPERRDDADNPLRVAFGEANDKIDERSRRFG-VQVEDGFSTENAVIVRE 306
 QY 485 TCG-----RGAVGRASWLYKHIFPFLIRYDVTGEPFRDQGHCMATSPGEPGLVAP 538
 DB 307 PGTPSGISGRGAHVAVVNGETVTECAVARFD-----AGALTNADDAIGELVN 355
 QY 539 VSGQSPFLGYAGPELAOGKLTMDVERPDPVFENCGDLLVCDGFLRFHRTGDFRFX 598
 DB 356 TTSGGFTGYNDPEANARM---RHG--MYWSLDLYRDSBGNTYIAGRTADMRVD 408
 QY 599 GENVATTEVAEVALDFLOEVNVYGVTPGHEGRAGNAALVLRPEHAL-DLMOLYTVH 656
 DB 409 GENLTAAPTRILLRKAINRVAVVP-DEYVGDQWAAVLVLRAGDTPDPAFEAFLDA 467
 QY 657 SENLPPYARPRFLQESLATTETTRQOQYRANESFDSITSLDPLVYLDQAVGAY 712
 DB 468 QPLSTKARPRYRIADLPSTATHKVLKROLIDEGTAVGK-ADTLWREPGSAY 522

Search completed: June 30, 2004, 17:36:06
 Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2004, 17:23:58 ; Search time 60 Seconds

(without alignments)
3437.661 Million cell updates/sec

Title: US-10-030-226-2

Perfect score: 3843

Sequence: 1 MGVCQTRPAPWKEKSQLERRA.....AYLPITARYSALLAGNLRI 730

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Database : A_Geneseq29Jan04:*
1: geneseqp1990s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3843	100.0	730	2	AAV41699 Human PRO
2	3843	100.0	730	3	AAAB44255 Human PRO
3	3843	100.0	730	3	AAAB24054 Human PRO
4	3843	100.0	730	4	AAAB60388 Human fat
5	3843	100.0	730	4	AAAB83779 Human FAT
6	3843	100.0	730	6	ABO25201 Novel hum
7	3843	100.0	730	6	ABU72207 Novel hum
8	3843	100.0	730	6	ADA84115 Human POM
9	3843	100.0	730	6	ABU84887 Human sec
10	3843	100.0	730	6	ABU61085 Human PRO
11	3843	100.0	730	6	ABU80354 Human sec
12	3843	100.0	730	6	ADA24641 Novel hum
13	3843	100.0	730	6	ABO19656 Novel hum
14	3843	100.0	730	6	ADA12302 Human sec
15	3843	100.0	730	6	ABO19547 Novel hum
16	3843	100.0	730	7	ADB73608 Human PRO
17	3843	100.0	730	7	ADB73624 Human PRO
18	3843	100.0	730	7	ADC43750 Human sec
19	3843	100.0	730	7	ADC61510 Human sec
20	3843	100.0	730	7	ADC63474 Human sec
21	3843	100.0	730	7	ADC66574 Human sec
22	3843	100.0	730	7	ADC68698 Human sec
23	3843	100.0	730	7	ADC62758 Human sec
24	3843	100.0	730	7	ADC67823 Human sec
25	3843	100.0	730	7	ADC41143 Human sec

26	3843	100.0	730	7	ADC67198 Human sec
27	3843	100.0	730	7	ADC62134 Human sec
28	3843	100.0	730	7	ADC41767 Human sec
29	3843	100.0	730	7	ADC49136 Human sec
30	3843	100.0	730	7	ADC35190 Human sec
31	3843	100.0	730	7	ADDE16304 Human sec
32	3843	100.0	730	7	ADDD72919 Human sec
33	3843	100.0	730	7	ADD72277 Human sec
34	3843	100.0	730	7	ADDE16928 Human sec
35	3843	100.0	730	8	ADDE48436 Human sec
36	3843	100.0	730	8	ADDE89537 Human sec
37	3843	100.0	811	6	AAE32069 Human TRI
38	3683	95.8	700	4	AAAB95135 Human PRO
39	3657	95.2	702	2	AAV14969 Amino ac1
40	3657	95.2	702	4	AAAB83278 Human FAT
41	3592	93.5	683	4	AAAG62172 Human gen
42	3525.5	91.7	758	6	AAE32083 Human TRI
43	3474	90.4	667	5	ABG63572 Human alD
44	2759	71.8	613	2	AAV14933 Amino ac1
45	2759	71.8	613	4	AAAB83271 Murine FA

ALIGNMENTS

RESULT 1	AAV41699	AAV41699 standard; protein; 730 AA.
ID	AAV41699	
XX	AAV41699;	
AC	AAV41699;	
DT	07-DEC-1999 (first entry)	
XX		
DE	Human PRO703 protein sequence.	
XX		
KW	Human; PRO; EST; expressed sequence tag; PCR primer; hybridisation;	
KM	probe; blood coagulation disorder; cancer; cellular adhesion disorder;	
XX	secreted protein; transmembrane protein.	
OS	Homo sapiens.	
XX		
PN	WO9946281-A2.	
PD	16-SEP-1999.	
XX		
PF	08-MAR-1999;	99WO-US005028.
XX		
PR	10-MAR-1998;	98US-0077450P.
PR	11-MAR-1998;	98US-0077632P.
PR	11-MAR-1998;	98US-0077641P.
PR	11-MAR-1998;	98US-0077649P.
PR	12-MAR-1998;	98US-0077791P.
PR	13-MAR-1998;	98US-0078004P.
PR	17-MAR-1998;	98US-00040220.
PR	17-MAR-1998;	98US-0078886P.
PR	20-MAR-1998;	98US-0078910P.
PR	20-MAR-1998;	98US-0078936P.
PR	20-MAR-1998;	98US-0078939P.
PR	25-MAR-1998;	98US-0079294P.
PR	26-MAR-1998;	98US-0079656P.
PR	27-MAR-1998;	98US-0079663P.
PR	27-MAR-1998;	98US-0079664P.
PR	27-MAR-1998;	98US-0079689P.
PR	27-MAR-1998;	98US-0079728P.
PR	27-MAR-1998;	98US-0079786P.
PR	30-MAR-1998;	98US-0079920P.
PR	30-MAR-1998;	98US-0079923P.
PR	31-MAR-1998;	98US-0080105P.
PR	31-MAR-1998;	98US-0080107P.
PR	31-MAR-1998;	98US-0080165P.
PR	31-MAR-1998;	98US-0080194P.
PR	01-APR-1998;	98US-0080327P.
PR	01-APR-1998;	98US-0080328P.

PR 01-APR-1998; 98US-0080333P.
 PR 01-APR-1998; 98US-0080334P.
 PR 08-APR-1998; 98US-0081049P.
 PR 08-APR-1998; 98US-0081070P.
 PR 08-APR-1998; 98US-0081071P.
 PR 09-APR-1998; 98US-0081195P.
 PR 09-APR-1998; 98US-0081203P.
 PR 09-APR-1998; 98US-0081229P.
 PR 15-APR-1998; 98US-0081817P.
 PR 15-APR-1998; 98US-0081838P.
 PR 15-APR-1998; 98US-0081952P.
 PR 15-APR-1998; 98US-0081955P.
 PR 21-APR-1998; 98US-0082568P.
 PR 21-APR-1998; 98US-0082569P.
 PR 22-APR-1998; 98US-0082700P.
 PR 22-APR-1998; 98US-0082704P.
 PR 23-APR-1998; 98US-0082804P.
 PR 23-APR-1998; 98US-0082867P.
 PR 23-APR-1998; 98US-0082966P.
 PR 27-APR-1998; 98US-0083336P.
 PR 28-APR-1998; 98US-0083322P.
 PR 29-APR-1998; 98US-0083392P.
 PR 29-APR-1998; 98US-0083495P.
 PR 29-APR-1998; 98US-0083496P.
 PR 29-APR-1998; 98US-0083499P.
 PR 29-APR-1998; 98US-0083500P.
 PR 29-APR-1998; 98US-0083545P.
 PR 29-APR-1998; 98US-0083549P.
 PR 29-APR-1998; 98US-0083558P.
 PR 29-APR-1998; 98US-0083559P.
 PR 30-APR-1998; 98US-0083742P.
 PR 05-MAY-1998; 98US-0084366P.
 PR 06-MAY-1998; 98US-0084414P.
 PR 06-MAY-1998; 98US-0084415P.
 PR 07-MAY-1998; 98US-0084598P.
 PR 07-MAY-1998; 98US-0084600P.
 PR 07-MAY-1998; 98US-0084627P.
 PR 07-MAY-1998; 98US-0084637P.
 PR 07-MAY-1998; 98US-0084639P.
 PR 07-MAY-1998; 98US-0084640P.
 PR 07-MAY-1998; 98US-0084643P.
 PR 13-MAY-1998; 98US-0085323P.
 PR 13-MAY-1998; 98US-0085338P.
 PR 13-MAY-1998; 98US-0085339P.
 PR 15-MAY-1998; 98US-0085573P.
 PR 15-MAY-1998; 98US-0085579P.
 PR 15-MAY-1998; 98US-0085580P.
 PR 15-MAY-1998; 98US-0085582P.
 PR 15-MAY-1998; 98US-0085589P.
 PR 15-MAY-1998; 98US-0085697P.
 PR 15-MAY-1998; 98US-0085700P.
 PR 15-MAY-1998; 98US-0085704P.
 PR 18-MAY-1998; 98US-0086023P.
 PR 22-MAY-1998; 98US-0086392P.
 PR 22-MAY-1998; 98US-0086414P.
 PR 22-MAY-1998; 98US-0086430P.
 PR 22-MAY-1998; 98US-0086486P.
 PR 28-MAY-1998; 98US-0087098P.
 PR 28-MAY-1998; 98US-0087106P.
 PR 28-MAY-1998; 98US-0087208P.
 PR 30-JUL-1998; 98US-0094651P.
 PR 11-SEP-1998; 98US-0100038P.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Wood WI, Goddard A, Gurney A, Yuan J, Baker KP, Chen J;
 XX
 DX WFI, 1999-551358/46.
 DR N-PSDB; AA233977.
 XX
 PT New secreted and transmembrane polypeptides and their polymuclectides,
 PT useful for treating blood coagulation disorders, cancers and cellular
 PT adhesion disorders.

XX
 PS Claim 12, Fig 39; 530pp; English.
 XX
 CC The present invention describes secreted and transmembrane polypeptides
 CC and their polymuclectides. The nucleotide sequences are useful as sources
 CC of probes, primers, for chromosome mapping, and for generation of
 CC antisense sequences. They can also be used to create transgenic animals.
 CC The proteins can be used to treat a variety of diseases and disorders,
 CC depending on their function. Diseases that may be treated include blood
 CC coagulation disorders, cancers and cellular adhesion disorders. They may
 CC also be used to raise antibodies. AA233891 to AA234338, and AA241685 to
 CC AA241774 represent polymuclectide and polypeptide sequence given in the
 CC exemplification of the present invention
 XX
 SQ Sequence 730 AA;
 Query Match 100.0%; Score 3843; DB 2; Length 730;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSVCQRTAPWKEKSKQLERAAAGFRGSGMFGSGMNVPIEEGSMALLPLPLLL 60
 DB 1 MGVCQRTAPWKEKSKQLERAAAGFRGSGMFGSGMNVPIEEGSMALLPLPLLL 60
 QY PLLLLKLHLPQULRWLPADLAPAVRALCCRALRAVALAAAAADPEGPGCSLAWRLAE 120
 DB PLLLLKLHLPQULRWLPADLAPAVRALCCRALRAVALAAAAADPEGPGCSLAWRLAE 120
 QY 61 LAQQRAAHTFLIGSRSPFSYSEKRESNRAAAFLEALCWDMDGPGDGSAGSGERA 180
 DB 61 LAQQRAAHTFLIGSRSPFSYSEKRESNRAAAFLEALCWDMDGPGDGSAGSGERA 180
 QY 121 LAQQRAAHTFLIGSRSPFSYSEKRESNRAAAFLEALCWDMDGPGDGSAGSGERA 180
 DB 121 LAQQRAAHTFLIGSRSPFSYSEKRESNRAAAFLEALCWDMDGPGDGSAGSGERA 180
 QY 181 APGAGDAAGSGAEFAGDGAAGGAAAPLSGATVALLPAGPEFLWFLGAKAGUR 240
 DB 181 APGAGDAAGSGAEFAGDGAAGGAAAPLSGATVALLPAGPEFLWFLGAKAGUR 240
 QY 241 TAFVPTALRRGPLHLCIRSCGARALVLAPEFLESELPDIPALRANGLHMAAGPGRHPAG 300
 DB 241 TAFVPTALRRGPLHLCIRSCGARALVLAPEFLESELPDIPALRANGLHMAAGPGRHPAG 300
 QY 301 ISDLLEVSAPVNDGVPVGVASSPQSTTDCVLYFTSGTGLPKAARISHKLIIQCGFYQ 360
 DB 301 ISDLLEVSAPVNDGVPVGVASSPQSTTDCVLYFTSGTGLPKAARISHKLIIQCGFYQ 360
 QY 361 LCGVHOEDVLYALPLTHMSGSLIGVCGMIGATVVLKSKTSAGQFWEDCOOHRVTYQ 420
 DB 361 LCGVHOEDVLYALPLTHMSGSLIGVCGMIGATVVLKSKTSAGQFWEDCOOHRVTYQ 420
 QY 421 YIGELCRYLYNPPSKAERGHKRLAVSGLRPDTERRFVRRRGPLQVLETGLTEGNYA 480
 DB 421 YIGELCRYLYNPPSKAERGHKRLAVSGLRPDTERRFVRRRGPLQVLETGLTEGNYA 480
 QY 481 TINYTGORGAVGRASWLYKHIPFSLIRYDVTGGEIRDPQGHCVATSPGEPGLVAPVS 540
 DB 481 TINYTGORGAVGRASWLYKHIPFSLIRYDVTGGEIRDPQGHCVATSPGEPGLVAPVS 540
 QY 541 QQSPPFLGAGGPELAGCKLKDVFRRGDDVFRTNGDILLVCDQGFLEFHRGTGTFWKKE 600
 DB 541 QQSPPFLGAGGPELAGCKLKDVFRRGDDVFRTNGDILLVCDQGFLEFHRGTGTFWKKE 600
 QY 601 NVATTEVAEVEFALDFEOEVNVVGVTVPGHEGAGNAALVREPPLADLMQVTHYSENL 660
 DB 601 NVATTEVAEVEFALDFEOEVNVVGVTVPGHEGAGNAALVREPPLADLMQVTHYSENL 660
 QY 661 PPYARPRFLRLOSATTEFFKQOKVPMANEGDPSLSDPLVYLDOAVGAYPLITARY 720
 DB 661 PPYARPRFLRLOSATTEFFKQOKVPMANEGDPSLSDPLVYLDOAVGAYPLITARY 720
 QY 721 SALLAGNLRI 730
 DB 721 SALLAGNLRI 730

RESULT 2
AAB44255 standard; protein; 730 AA.
AC AAB44255;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human PRO703 (UNQ367) protein sequence SEQ ID NO:102.
XX
KW Human; secreted protein; transmembrane protein; PRO; EST; cytosolic;
XX expressed sequence tag; detection; cancer.
OS Homo sapiens.
XX
EN WO200053756-A2.
XX
PD 14-SEP-2000.
XX
PE 18-FEB-2000; 2000WO-US004341.
XX
PR 08-MAR-1999; 99WO-US005028.
XX 12-MAR-1999; 99US-0123957P.
XX 29-MAR-1999; 99US-0126773P.
XX 21-APR-1999; 99US-0130233P.
XX 28-APR-1999; 99US-0131445P.
XX 14-MAY-1999; 99US-0134287P.
XX 23-JUN-1999; 99US-0141037P.
XX 26-JUL-1999; 99US-0145698P.
XX 29-OCT-1999; 99US-0162506P.
XX 30-NOV-1999; 99WO-US028313.
XX 02-DEC-1999; 99WO-US028551.
XX 16-DEC-1999; 99WO-US031025.
XX 30-DEC-1999; 99WO-US031243.
XX 30-DEC-1999; 99WO-US031274.
XX 05-JAN-2000; 2000WO-US000219.
XX 06-JAN-2000; 2000WO-US000277.
XX 06-JAN-2000; 2000WO-US000376.
XX
PA (GENETH) GENENTECH INC.
XX
PI Ashkenazi AJ, Baker KP, Bolstein D, Desnoyers L, Eaton DL;
PI Ferrare N, Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME;
PI Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan MJ;
PI Kljavin IJ, Kuo SS, Napier MA, Pan J, Paoni NF, Roy MA, Shelton DL;
PI Stewart TX, Tumas D, Williams EM, Wood WI;
XX
DR WPI; 2000-611443/58.
XX N-PSDB; AACT8481.
XX
PT Novel PRO polypeptides and polynucleotides used in detection methods, to
PT target bioactive molecules to specific cells, and to modulate cellular
PT activities.
XX
PS Claim 12, Fig 39; 636p; English.
XX
XX AACT8458 to AACT8599 represent polynucleotide and EST (expressed sequence
XX tag) sequences which encode secreted or transmembrane PRO polypeptides.
XX The PRO polynucleotides and polypeptides have cytosolic activity. The
XX polynucleotides and polypeptides can be used for detecting the presence
XX of PRO polypeptides in samples, for linking bioactive molecules to cells
XX and for modulating biological activities of cells, using the polypeptides
XX for specific targeting. The polypeptide targeting can be used to kill the
XX target cells, e.g. for the treatment of cancers. The polypeptide pairs
XX provide specific targeting of bioactive molecules to cells. AACT8600 to
XX AACT8987 represent PCR primers and probes used in the isolation of the
XX PRO polynucleotide sequences
XX
XX Sequence 730 AA;
XX
XX Query Match 100.0%; Score 3843; DB 3; Length 730;
XX Best Local Similarity 100.0%; Pred. No. 0;

Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGVCORTBAPWKEKSEQLERALLGFRKGGSGMFAFGNNOVPIREAGSMALLPLPLLL 60
Db 1 MGVCORTBAPWKEKSEQLERALLGFRKGGSGMFAFGNNOVPIREAGSMALLPLPLLL 60
QY PLILLKILHMPQILRWLPADLAFARALCCRRALRABALAAAAADPEPGGCSLAWRLAE 120
Db PLILLKILHMPQILRWLPADLAFARALCCRRALRABALAAAAADPEPGGCSLAWRLAE 120
QY 121 LAQORAAHFTLHGSRFRFSYSEARESNRAARFALMDMDPDGDSGSGAGGCEAA 180
Db 121 LAQORAAHFTLHGSRFRFSYSEARESNRAARFALMDMDPDGDSGSGAGGCEAA 180
QY 181 APGAGDAAAGSAGAEFAGDGAARGGGAAPLSFGATVALLPAGPEFLMWGLAKAGLR 240
Db 181 APGAGDAAAGSAGAEFAGDGAARGGGAAPLSFGATVALLPAGPEFLMWGLAKAGLR 240
QY 241 TAFVPTALRRGPLHCLRSCGARALVLAPEFLSELEPDLPALRAMELHMAAGPGTHPAG 300
Db 241 TAFVPTALRRGPLHCLRSCGARALVLAPEFLSELEPDLPALRAMELHMAAGPGTHPAG 300
QY 301 ISDLAEVSAEVDGVPYGLTSSPOSTITDCLYITFTSGTGLPKAAARISHLKLQCGFYQ 360
Db 301 ISDLAEVSAEVDGVPYGLTSSPOSTITDCLYITFTSGTGLPKAAARISHLKLQCGFYQ 360
QY 361 LCGVHOEDVIYIALPLYNHSGSLGIVGCMGIGATVILKSKFSAGQFWEDCCQHRVTYQ 420
Db 361 LCGVHOEDVIYIALPLYNHSGSLGIVGCMGIGATVILKSKFSAGQFWEDCCQHRVTYQ 420
QY 421 YIGELCRYLVNPPPSAERGHKRYRLAVSGGLPDMTERVRRFPGPOVETGLGEGNVA 480
Db 421 YIGELCRYLVNPPPSAERGHKRYRLAVSGGLPDMTERVRRFPGPOVETGLGEGNVA 480
QY 481 TIINYTGQKAVGASNLVYKHFPPSLIRYDVTGSEIRDPQHCNATSPGEGELVAPVS 540
Db 481 TIINYTGQKAVGASNLVYKHFPPSLIRYDVTGSEIRDPQHCNATSPGEGELVAPVS 540
QY 541 QOSPFGLVAGGPELQAGKLLKDFRPGDVFYFNTGDLVDDGFLRFHRTDPTRWKKE 600
Db 541 QOSPFGLVAGGPELQAGKLLKDFRPGDVFYFNTGDLVDDGFLRFHRTDPTRWKKE 600
QY 601 NVATTEVAEVEFALPFIQEVNYYGVTPGHEGRAGAAALVLRPPHALDLMQYTHVSENL 660
Db 601 NVATTEVAEVEFALPFIQEVNYYGVTPGHEGRAGAAALVLRPPHALDLMQYTHVSENL 660
QY 661 PPAARPRFLRDESLATTETTFKQKVRANESGPDSTLSDPLVILDQAVGAYLPLTTARY 720
Db 661 PPAARPRFLRDESLATTETTFKQKVRANESGPDSTLSDPLVILDQAVGAYLPLTTARY 720
QY 721 SALLAGNLRI 730
Db 721 SALLAGNLRI 730
RESULT 3
AAB24054
ID AAB24054 standard; protein; 730 AA.
XX
AC AAB24054;
XX
DT 25-JAN-2001 (first entry)
XX
DE Human PRO703 protein sequence SEQ ID NO:29.
XX
KW Human; tumor; diagnosis; neoplastic disease; identification; cancer;
KW tumorigenesis; detection; neoplastic cell growth; proliferation;
KW cytostatic; antiinflammatory; immunomodulatory; inflammatory disorder;
KW immunological disorder.
XX
OS Homo sapiens.
XX
XX WO200053754-A1.

XX 14-SEP-2000. ✓
 PD
 XX
 PF 06-JAN-2000; 2000MO-US000277.
 XX
 PR 08-MAR-1999; 99MO-US005028.
 XX
 PR 12-MAR-1999; 99US-0123957P.
 PR 29-MAR-1999; 99US-012673P.
 PR 21-APR-1999; 99US-013023P.
 PR 28-APR-1999; 99US-0131445P.
 PR 05-OCT-1999; 99MO-US023089.
 PR 30-NOV-1999; 99MO-US028313.
 PR 02-DEC-1999; 99MO-US028551.
 PR 30-DEC-1999; 99MO-US028564.
 PR 30-DEC-1999; 99MO-US031243.
 PR 30-DEC-1999; 99MO-US031274.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Baker KP, Desauvage FJ, Goddard A, Gurney AL, Klein RD, Roy MA,
 PI Wood WJ;
 XX
 DR WPI: 2000-572269/53.
 DR N-PSDB; AAC58239.
 XX
 XX New isolated antibody for use in compositions and methods for the
 PT diagnosis and treatment of neoplastic cell growth and proliferation in
 PT mammals, including humans, and in monitoring tumor treatment.
 XX
 PS Claim 61; Fig 29; 1999p; English.
 CC The present invention describes an isolated antibody (Ab) that binds to
 CC one of the human proteins (p) designated PRO213, PRO1330, PRO1449,
 CC PRO324, PRO351, PRO362, PRO615, PRO538, PRO3664, PRO618,
 CC PRO772, PRO703, PRO792 or PRO474. The Ab can be used in compositions and
 CC methods for the diagnosis and treatment of neoplastic cell growth and
 CC proliferation in mammals, including humans. Genes and polypeptides
 CC encoded by them, that are amplified in the genome of a tumour cell, can
 CC be identified and are useful targets for the treatment and prevention of
 CC certain cancers and may be used to monitor tumour treatment. Compounds
 CC that inhibit the expression or activity of the identified polypeptides
 CC can be identified and used as antagonists. Benign or malignant tumours,
 CC inflammatory disorders and immunological disorders can be treated.
 CC AAC58123 to AAC58224 represent hybridisation probes and PCR primers used
 CC in the isolation of the human PRO sequences. AAC58225 to AAC58241 and
 CC AAB24041 to AAB24056 represent human PRO polynucleotide and protein
 CC sequences given in the exemplification of the present invention
 CC
 XX
 SQ Sequence 730 AA;
 Query Match 100.0%; Score 3843; DB 3; Length 730;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGVCCRTAPWKEKQLEERAAAGFRKSGSGMFGASGNNQTVPIEAGSMAALLLPLILL 60
 DB 1 MGVCCRTAPWKEKQLEERAAAGFRKSGSGMFGASGNNQTVPIEAGSMAALLLPLILL 60
 QY 61 PLLILKTHMPLQRLMPLADIAFAVRALCKKRLRRALAAAADGEGGCGSLAMRLAE 120
 DB 61 PLLILKTHMPLQRLMPLADIAFAVRALCKKRLRRALAAAADGEGGCGSLAMRLAE 120
 QY 121 LAQORAAATFTLTHGSRFSYSSEERESNRAAFALRGMDWDPDGDGSGEGSAGEGERA 180
 DB 121 LAQORAAATFTLTHGSRFSYSSEERESNRAAFALRGMDWDPDGDGSGEGSAGEGERA 180
 QY 181 APAGGAAAAGSGAFEGSGDGAARGGGAAAPLSPGATVALLPAGGEPFLMWGLAKGLR 240
 DB 181 APAGGAAAAGSGAFEGSGDGAARGGGAAAPLSPGATVALLPAGGEPFLMWGLAKGLR 240
 QY 241 TAFVPTALRRGPLALHCLRS CGARALVLABEFLESIEPDLPALRANGHLHMAAGPTHPAG 300
 DB 241 TAFVPTALRRGPLALHCLRS CGARALVLABEFLESIEPDLPALRANGHLHMAAGPTHPAG 300

QY 301 ISDLIAEVSAAEDGVPVGYLSSPQSIDTICLYIFTS GTTGTPKKAIRISHLKLQCGPYQ 360
 DB 301 ISDLIAEVSAAEDGVPVGYLSSPQSIDTICLYIFTS GTTGTPKKAIRISHLKLQCGPYQ 360
 QY 361 LCGVHOEDVITYALPLTHMSGSLIGVGMGIGATVVLKSKPSAGCFMEDCOQHRTVFO 420
 DB 361 LCGVHOEDVITYALPLTHMSGSLIGVGMGIGATVVLKSKPSAGCFMEDCOQHRTVFO 420
 QY 421 YIGELCRYLVNQPSPKAERGHKVRLLAVSGRLPDTWERFVRFGFLQVLETYGLTEGVA 480
 DB 421 YIGELCRYLVNQPSPKAERGHKVRLLAVSGRLPDTWERFVRFGFLQVLETYGLTEGVA 480
 QY 481 TINTYGOARAVRASGLYHITPPSLIRVDVTGTPIDPOCHMATRSGRGLIVAVS 540
 DB 481 TINTYGOARAVRASGLYHITPPSLIRVDVTGTPIDPOCHMATRSGRGLIVAVS 540
 QY 541 QQSPFLGYAGPELAQGLKLDVFRPGDVFNTGDLVCDGGLRPHDRGTFRMKGE 600
 DB 541 QQSPFLGYAGPELAQGLKLDVFRPGDVFNTGDLVCDGGLRPHDRGTFRMKGE 600
 QY 601 NVATTEVAEVEFALDPLQEVNNTYGVTPGHEGRAGAAALVLRPPHALDMQLYTHVSENL 660
 DB 601 NVATTEVAEVEFALDPLQEVNNTYGVTPGHEGRAGAAALVLRPPHALDMQLYTHVSENL 660
 QY 661 PPYARPRFLRQESLATTETFPKQKVRMANEGFDPSLSDPYVVDQAVGAYLPITARY 720
 DB 661 PPYARPRFLRQESLATTETFPKQKVRMANEGFDPSLSDPYVVDQAVGAYLPITARY 720
 QY 721 SALLAGNLRI 730
 DB 721 SALLAGNLRI 730
 RESULT 4
 AAB60388
 ID AAB60388 standard; protein; 730 AA.
 AC AAB60388;
 XX
 DT 24-APR-2001 (first entry)
 XX
 DE Human fatty acid transporter PSEC67.
 XX
 KW Human; fatty acid transporter; PSEC67; long-chain fatty acid uptake;
 KW oleic acid; drug screening; gene therapy; metabolic disorder;
 KW cardiomyopathy; skeletal muscle disorders; renal failure.
 XX
 OS Homo sapiens.
 XX
 PN WO200104301-A1.
 XX
 PD 18-JAN-2001/
 XX
 PF 07-JUL-2000; 2000MO-US0004549.
 XX
 PR 08-JUL-1999; 99US-00194179.
 PR 18-OCT-1999; 99US-0159586P.
 PR 25-APR-2000; 2000JP-00128993.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Morikawa N, Masuno Y, Ota T, Isogai T, Nishikawa T, Kawai Y;
 PI WPI: 2001-138349/14.
 DR N-PSDB; AAF27417.
 XX
 PT Fatty acid transporter protein and encoded gene PSEC67 cloned from human
 PT cDNA library, with activity of oleic acid incorporation, useful as target
 PT molecule of preventives or remedies of fatty-acid metabolic disorders.
 PS Claim 1; Page 48-51; 58pp; Japanese.

CC The invention relates to a novel human fatty acid transporter, PSEC67
 CC (AAB60388), and to cDNA encoding it (AA27417). PSEC67 is responsible for
 CC the uptake of oleic acid into cells. The invention also relates to
 CC vectors and host cells comprising a PSEC67 nucleic acid; the recombinant
 CC production of PSEC67; an antibody against PSEC67; methods of screening
 CC for compounds which can regulate the uptake of long-chain fatty acids
 CC into cells; and the compounds thus identified. The PSEC67 protein and the
 CC gene encoding it are useful as targets for the treatment or prevention of
 CC diseases associated with defective metabolism of long-chain fatty acids.
 CC Such diseases include cardiomyopathy, skeletal muscle disorders, or renal
 CC failure. PSEC67 nucleic acids may also be used in gene therapy for such
 CC disorders. The present sequence represents PSEC67

XX
 CC Sequence 730 AA:

Query Match 100.0%; Score 3843; DB 4; Length 730;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVCORTRAPWKEKSOLEERALGFRKSGSGMFASGNNQTVPIEAGSMALLLLPLILL 60
 DB 1 MGVCORTRAPWKEKSOLEERALGFRKSGSGMFASGNNQTVPIEAGSMALLLLPLILL 60
 QY 61 PLILKLHLMPQLRWLPADLAFAVRALCKKRALRAALAAAADPEGEGCSLAWRLAE 120
 DB 61 PLILKLHLMPQLRWLPADLAFAVRALCKKRALRAALAAAADPEGEGCSLAWRLAE 120
 QY 121 LAOQRAAHTFLIHGSRFRSYSEARESNRAARAFRLALGMDWGPDDGSGSGSAGEERA 180
 DB 121 LAOQRAAHTFLIHGSRFRSYSEARESNRAARAFRLALGMDWGPDDGSGSGSAGEERA 180
 QY 121 LAOQRAAHTFLIHGSRFRSYSEARESNRAARAFRLALGMDWGPDDGSGSGSAGEERA 180
 DB 121 LAOQRAAHTFLIHGSRFRSYSEARESNRAARAFRLALGMDWGPDDGSGSGSAGEERA 180
 QY 181 APGAGDAAGSAGAEFAGDGAARGGGAAAPLSPGATVALLPAPPEFLMWFGIAKXGLR 240
 DB 181 APGAGDAAGSAGAEFAGDGAARGGGAAAPLSPGATVALLPAPPEFLMWFGIAKXGLR 240
 QY 241 TAFVPTALRRGPPLHCLRSAGARALVLAPEFLSLEPDLPALRMGLHMAAGFTHPAG 300
 DB 241 TAFVPTALRRGPPLHCLRSAGARALVLAPEFLSLEPDLPALRMGLHMAAGFTHPAG 300
 QY 301 ISDLAEVSAEVDGVPGLTSSPISITDTCLYFTSGTGLPKARISHKXITLCCQGFYQ 360
 DB 301 ISDLAEVSAEVDGVPGLTSSPISITDTCLYFTSGTGLPKARISHKXITLCCQGFYQ 360
 QY 361 LCGVHOEDVIYALPLVYHMSGSLIGIVCGMIGATVVLKSKFSAGQFWEDECQHRVTVFQ 420
 DB 361 LCGVHOEDVIYALPLVYHMSGSLIGIVCGMIGATVVLKSKFSAGQFWEDECQHRVTVFQ 420
 QY 421 YIGELCRVLVNOPPSKARGHKVRLAVSGSLRPTWTFVRRFPLQVLEFYGLTEENVA 480
 DB 421 YIGELCRVLVNOPPSKARGHKVRLAVSGSLRPTWTFVRRFPLQVLEFYGLTEENVA 480
 QY 481 TINTYGOGAGVGRASMLYKHIFPESLIRYDVTGTGPRIIDPGHKCAISPGSGILVAPVS 540
 DB 481 TINTYGOGAGVGRASMLYKHIFPESLIRYDVTGTGPRIIDPGHKCAISPGSGILVAPVS 540
 QY 541 QOSPFGLYAGGBELAQGLKLDVFRPGDVFFNTGDLVCDQGFRLFRDRTGDFPRNKG 600
 DB 541 QOSPFGLYAGGBELAQGLKLDVFRPGDVFFNTGDLVCDQGFRLFRDRTGDFPRNKG 600
 QY 601 NVATTEVEVEFEALDFLOENVVYGVTVGHSRAGMALVLRPPALDLMGLYHVSNTL 660
 DB 601 NVATTEVEVEFEALDFLOENVVYGVTVGHSRAGMALVLRPPALDLMGLYHVSNTL 660
 QY 661 PEYARPRFLQESLATTETEFKQKVRMANEGFDSTLSDPLVYLDQAVGAYLPLTARY 720
 DB 661 PEYARPRFLQESLATTETEFKQKVRMANEGFDSTLSDPLVYLDQAVGAYLPLTARY 720
 QY 721 SALLAGNLRI 730
 DB 721 SALLAGNLRI 730

RESULT 5

AA83279
 ID AAB83279 standard; protein; 730 AA.
 XX
 AC AAB83279;
 XX
 DT 06-JUL-2001 (first entry)
 XX
 DE Human FATP3 SEQ ID NO: 117.
 XX
 XX Fatty acid transport protein; FATP, human; mouse; rat; rice blast fungus;
 KM yeast; fat absorption; obesity; diabetes; heart disease; hyperlipidemia;
 KM weight control; tuberculosis; TB; anti-fungal.
 XX
 OS Homo sapiens.
 XX
 PN WO200121795-A2.
 XX
 PD 29-MAR-2001.
 XX
 PF 21-SEP-2000; 2000MO-US025891.
 XX
 PR 23-SEP-1999; 99US-00405504.
 PR 23-SEP-1999; 99US-00405505.
 PR 16-DEC-1999; 99US-00465280.
 PR 17-FEB-2000; 2000US-00506252.
 PR 06-JUL-2000; 2000US-00611197.
 XX
 PA (MHED) WHITEHEAD INST BIOMEDICAL RES.
 PA (MILL) MILLIENIUM PHARM INC.
 XX
 PI Stahl A, Hirsch DJ, Lodish HF, Gimeno RE, Tartaglia LA;
 DR WPI; 2001-354783/37.
 DR N-PSDB; AAF89054.
 XX
 PT New fatty acid transport proteins (FATPs) useful for the manufacture of
 PT medicament for treating obesity, diabetes and heart disease.
 XX
 PS Claim 31; Fig 112; 287bp; English.
 XX
 CC The present invention provides the protein and coding sequences of fatty
 CC acid transport proteins (FATPs) from a number of species, including
 CC FATP1, FATP2, FATP3, FATP4, FATP5 and FATP6 from the human, FATP1-FATP5
 CC from the mouse, FATP4 and b from C. elegans, and FATP from Aspergillus
 CC nidulans, Drosophila, zebrafish, Magnaporthe grisea, Mycobacterium
 CC tuberculosis and Cochliobolus heterostrophus. The FATP from M.
 CC tuberculosis can be used to identify inhibitors which can then be used to
 CC treat TB. That from M. grisea (also known as rice blast fungus) can be
 CC used to develop anti-fungal agents capable of preventing infection of
 CC rice. Those from the human can be used to develop treatments for
 CC diabetes, heart disease, obesity, hyperlipidemia and weight control. The
 CC present sequence is one of the sequences described in the exemplification
 CC of the invention
 XX
 SQ Sequence 730 AA:
 Query Match 100.0%; Score 3843; DB 4; Length 730;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVCORTRAPWKEKSOLEERALGFRKSGSGMFASGNNQTVPIEAGSMALLLLPLILL 60
 DB 1 MGVCORTRAPWKEKSOLEERALGFRKSGSGMFASGNNQTVPIEAGSMALLLLPLILL 60
 QY 61 PLILKLHLMPQLRWLPADLAFAVRALCKKRALRAALAAAADPEGEGCSLAWRLAE 120
 DB 61 PLILKLHLMPQLRWLPADLAFAVRALCKKRALRAALAAAADPEGEGCSLAWRLAE 120
 QY 121 LAOQRAAHTFLIHGSRFRSYSEARESNRAARAFRLALGMDWGPDDGSGSGSAGEERA 180
 DB 121 LAOQRAAHTFLIHGSRFRSYSEARESNRAARAFRLALGMDWGPDDGSGSGSAGEERA 180
 QY 181 APGAGDAAGSAGAEFAGDGAARGGGAAAPLSPGATVALLPAPPEFLMWFGIAKXGLR 240

Db 181 ABOGDAAGSGAERFAGGGAARGGAAAPLSPGATVALLPAGPEFLMLWGLAKAELR 240
 Qy 241 TAFVPTALRGPELHLCLRSQCARALVLAPEFLESLEPDLPALRAMGLHMAAGPOTHPAG 300
 Db 244 TAFVPTALRGPELHLCLRSQCARALVLAPEFLESLEPDLPALRAMGLHMAAGPOTHPAG 300
 Qy 301 ISDLAIVSAEVDGPPPGYLLSSPQITDTCLYIFTSSTTGLPKARISHKLILQCGFYQ 360
 Db 301 ISDLAIVSAEVDGPPPGYLLSSPQITDTCLYIFTSSTTGLPKARISHKLILQCGFYQ 360
 Qy 361 LCGVHOEDVIYIALPLHYMSGSLIGVCGMGIGATVVKSKFSAGQFQFEDCQHRVTYFQ 420
 Db 361 LCGVHOEDVIYIALPLHYMSGSLIGVCGMGIGATVVKSKFSAGQFQFEDCQHRVTYFQ 420
 Qy 421 YIGELCRILYVNOPPSKAEKRGKRLAVSGSLRPDTEWFRFRFGPLQYLETYGLTEGVA 480
 Db 421 YIGELCRILYVNOPPSKAEKRGKRLAVSGSLRPDTEWFRFRFGPLQYLETYGLTEGVA 480
 Qy 481 TINYTGORGAVRAMVLYKHIFPFSLIRDYDTGEPIDPOGHCAATSPGEPGLLVAVS 540
 Db 481 TINYTGORGAVRAMVLYKHIFPFSLIRDYDTGEPIDPOGHCAATSPGEPGLLVAVS 540
 Qy 541 QCSPLGYAGGPBELAQGLKLDKDFRPGDVFPNTGDLVCDQGFLEFHDRTGDTFRWKE 600
 Db 541 QCSPLGYAGGPBELAQGLKLDKDFRPGDVFPNTGDLVCDQGFLEFHDRTGDTFRWKE 600
 Qy 601 NVATTEVEVEFPEALPFLQEVNYYGVTVPGHSGRAGMALVLRPPHALDMLQTYHVSNTL 660
 Db 601 NVATTEVEVEFPEALPFLQEVNYYGVTVPGHSGRAGMALVLRPPHALDMLQTYHVSNTL 660
 Qy 661 PPYAPRPRFLRQESLATTETFKQOKVRMANEGFDBSTLSDPLVYLDAVGAYLPLTARY 720
 Db 661 PPYAPRPRFLRQESLATTETFKQOKVRMANEGFDBSTLSDPLVYLDAVGAYLPLTARY 720
 Qy 721 SALLAGNLR 730
 Db 721 SALLAGNLR 730

RESULT 6
 ABO25201
 ID ABO25201 standard; protein; 730 AA.
 XX ABO25201;
 AC ABO25201;
 XX ABO25201;
 DT 09-SEP-2003 (first entry)
 XX
 DE Novel human secreted and transmembrane protein PRO703.
 XX
 KM Human; secreted and transmembrane protein; PRO; virucide; gene therapy;
 KM cell death; growth induction cascade; blood coagulation cascade;
 KM viral infection.
 XX
 OS Homo sapiens.
 XX
 PN US2003050239-A1.
 XX
 PD 13-MAR-2003.
 XX
 PF 15-OCT-2001; 2001US-00978191.
 XX
 XX 17-OCT-1997; 97US-0062250P.
 PR 03-NOV-1997; 97US-0064248P.
 PR 13-NOV-1997; 97US-0065311P.
 PR 21-NOV-1997; 97US-0066364P.
 PR 10-MAR-1998; 98US-0077450P.
 PR 11-MAR-1998; 98US-0077632P.
 PR 11-MAR-1998; 98US-0077641P.
 PR 11-MAR-1998; 98US-0077649P.
 PR 12-MAR-1998; 98US-0077791P.
 PR 13-MAR-1998; 98US-0078004P.
 PR 17-MAR-1998; 98US-00040220.
 PR 17-MAR-1998; 98US-00040220.

PR 20-MAR-1998; 98US-0078886P.
 PR 20-MAR-1998; 98US-0078910P.
 PR 20-MAR-1998; 98US-0078936P.
 PR 20-MAR-1998; 98US-0078939P.
 PR 25-MAR-1998; 98US-0079294P.
 PR 26-MAR-1998; 98US-0079656P.
 PR 27-MAR-1998; 98US-0079663P.
 PR 27-MAR-1998; 98US-0079664P.
 PR 27-MAR-1998; 98US-0079689P.
 PR 27-MAR-1998; 98US-0079728P.
 PR 30-MAR-1998; 98US-0079786P.
 PR 30-MAR-1998; 98US-0079920P.
 PR 30-MAR-1998; 98US-0079923P.
 PR 31-MAR-1998; 98US-0080105P.
 PR 31-MAR-1998; 98US-0080107P.
 PR 31-MAR-1998; 98US-0080165P.
 PR 31-MAR-1998; 98US-0080194P.
 PR 01-APR-1998; 98US-0080328P.
 PR 01-APR-1998; 98US-0080328P.
 PR 01-APR-1998; 98US-0080334P.
 PR 08-APR-1998; 98US-0081049P.
 PR 08-APR-1998; 98US-0081070P.
 PR 08-APR-1998; 98US-0081071P.
 PR 09-APR-1998; 98US-0081195P.
 PR 09-APR-1998; 98US-0081203P.
 PR 09-APR-1998; 98US-0081229P.
 PR 15-APR-1998; 98US-0081817P.
 PR 15-APR-1998; 98US-0081819P.
 PR 15-APR-1998; 98US-0081838P.
 PR 15-APR-1998; 98US-0081952P.
 PR 15-APR-1998; 98US-0081955P.
 PR 21-APR-1998; 98US-0082688P.
 PR 21-APR-1998; 98US-0082689P.
 PR 22-APR-1998; 98US-0082700P.
 PR 22-APR-1998; 98US-0082704P.
 PR 22-APR-1998; 98US-0082797P.
 PR 22-APR-1998; 98US-0082804P.
 PR 23-APR-1998; 98US-0082796P.
 PR 27-APR-1998; 98US-0083336P.
 PR 28-APR-1998; 98US-0083322P.
 PR 29-APR-1998; 98US-0083392P.
 PR 29-APR-1998; 98US-0083495P.
 PR 29-APR-1998; 98US-0083496P.
 PR 29-APR-1998; 98US-0083499P.
 PR 29-APR-1998; 98US-0083500P.
 PR 29-APR-1998; 98US-0083545P.
 PR 29-APR-1998; 98US-0083545P.
 PR 29-APR-1998; 98US-0083558P.
 PR 29-APR-1998; 98US-0083558P.
 PR 30-APR-1998; 98US-0083559P.
 PR 30-APR-1998; 98US-0083742P.
 PR 05-MAY-1998; 98US-0084366P.
 PR 06-MAY-1998; 98US-0084414P.
 PR 06-MAY-1998; 98US-0084414P.
 PR 07-MAY-1998; 98US-0084589P.
 PR 07-MAY-1998; 98US-0084589P.
 PR 07-MAY-1998; 98US-0084600P.
 PR 07-MAY-1998; 98US-0084627P.
 PR 07-MAY-1998; 98US-0084637P.
 PR 07-MAY-1998; 98US-0084637P.
 PR 07-MAY-1998; 98US-0084639P.
 PR 07-MAY-1998; 98US-0084640P.
 PR 07-MAY-1998; 98US-0084643P.
 PR 13-MAY-1998; 98US-0085323P.
 PR 13-MAY-1998; 98US-0085338P.
 PR 13-MAY-1998; 98US-0085339P.
 PR 15-MAY-1998; 98US-0085573P.
 PR 15-MAY-1998; 98US-0085579P.
 PR 15-MAY-1998; 98US-0085580P.
 PR 15-MAY-1998; 98US-0085582P.
 PR 15-MAY-1998; 98US-0085689P.
 PR 15-MAY-1998; 98US-0085697P.
 PR 15-MAY-1998; 98US-0085700P.
 PR 15-MAY-1998; 98US-0085704P.
 PR 18-MAY-1998; 98US-00856023P.

```
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086414P.
PR 22-MAY-1998; 98US-0086430P.
PR 22-MAY-1998; 98US-0086466P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087106P.
PR 28-MAY-1998; 98US-0087208P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090863P.
PR 01-JUL-1998; 98US-0091359P.
PR 30-JUL-1998; 98US-0094651P.
PR 11-SEP-1998; 98US-0100038P.
PR 07-OCT-1998; 98US-00168978.
PR 07-OCT-1998; 98US-0021141.
PR 02-NOV-1998; 98US-00184216.
PR 06-NOV-1998; 98US-00187368.
PR 20-NOV-1998; 98US-0109304P.
PR 20-NOV-1998; 98US-00202054.
PR 22-DEC-1998; 98US-00218517.
PR 22-DEC-1998; 98US-0113296P.
PR 23-DEC-1998; 98US-0113621P.
PR 05-JAN-1999; 99US-00254465.
PR 08-MAR-1999; 99US-00254465.
PR 10-MAR-1999; 99US-00265686.
PR 10-MAR-1999; 99US-00265686.
PR 12-MAR-1999; 99US-00267213.
PR 12-MAR-1999; 99US-0123957P.
PR 29-MAR-1999; 99US-0126773P.
PR 12-APR-1999; 99US-00284291.
PR 21-APR-1999; 99US-0130232P.
PR 26-APR-1999; 99US-0131022P.
PR 28-APR-1999; 99US-0131445P.
PR 14-MAY-1999; 99US-00311832.
PR 14-MAY-1999; 99US-0134287P.
PR 14-MAY-1999; 99US-0031073.
PR 02-JUN-1999; 99US-0031252.
PR 16-JUN-1999; 99US-0139557P.
PR 23-JUN-1999; 99US-0141037P.
PR 07-JUL-1999; 99US-0142680P.
PR 26-JUL-1999; 99US-0145698P.
PR 28-JUL-1999; 99US-0146222P.
PR 25-AUG-1999; 99US-00380137.
PR 25-AUG-1999; 99US-00380138.
PR 25-AUG-1999; 99US-00380142.
PR 29-OCT-1999; 99US-0162506P.
PR 30-NOV-1999; 99US-0028313.
PR 02-DEC-1999; 99US-0028551.
PR 02-DEC-1999; 99US-0028565.
PR 16-DEC-1999; 99US-0031254.
PR 30-DEC-1999; 99US-0031274.
PR 03-JAN-2000; 2000US-0000219.
PR 06-JAN-2000; 2000US-0000219.
PR 06-JAN-2000; 2000US-0000376.
PR 11-FEB-2000; 2000US-0003565.
PR 18-FEB-2000; 2000US-0004341.
PR 24-FEB-2000; 2000US-0005004.
PR 02-MAR-2000; 2000US-0005841.
PR 10-MAR-2000; 2000US-0006319.
PR 21-MAR-2000; 2000US-0007532.
PR 30-MAR-2000; 2000US-0008439.
PR 17-MAY-2000; 2000US-0013705.
PR 22-MAY-2000; 2000US-0014042.
PR 30-MAY-2000; 2000US-0014941.
PR 02-JUN-2000; 2000US-0015264.
PR 28-JUL-2000; 2000US-0020710.
PR 24-AUG-2000; 2000US-0023338.
PR 08-NOV-2000; 2000US-00709238.
PR 27-NOV-2000; 2000US-00723749.
PR 01-DEC-2000; 2000US-0072678.

PR 20-DEC-2000; 2000US-00747255.
PR 20-DEC-2000; 2000US-0074956.
PR 28-FEB-2001; 2001US-0006520.
PR 28-MAR-2001; 2001US-00816744.
PR 22-MAR-2001; 2001US-00816920.
PR 22-MAR-2001; 2001US-00816920.
PR 22-MAR-2001; 2001US-00816920.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 25-MAY-2001; 2001US-00854280.
PR 01-JUN-2001; 2001US-00872035.
PR 01-JUN-2001; 2001US-00872035.
PR 05-JUN-2001; 2001US-00874503.
PR 14-JUN-2001; 2001US-00882636.
PR 19-JUN-2001; 2001US-00886342.
PR 20-JUN-2001; 2001US-00886342.
PR 20-JUN-2001; 2001US-00886342.
PR 29-JUL-2001; 2001US-00921735.
PR 30-JUL-2001; 2001US-00921735.

( GETH ) GENENTECH INC.
PR Ashkenazi A, Baker KP, Botstein D, Desnovers J, Eaton DJ,
PI Ferrara N, Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME,
Query Match 100.0%; Score 3843; DB 6; Length 730;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVCQRTAPWKEKSGLEBALGFRKSGSGMPASGNNQVPIEBAGSMALLPLLL 60
D 1 MGVCQRTAPWKEKSGLEBALGFRKSGSGMPASGNNQVPIEBAGSMALLPLLL 60
QY 61 PLILKTLHMPQRLPADLAFAYRALCKKRLARALAAAADPREGGSLMRLAE 120
D 61 PLILKTLHMPQRLPADLAFAYRALCKKRLARALAAAADPREGGSLMRLAE 120
QY 121 LAQRAAHTFLIHGSRFYSSEARESNRARAFLALGMDGPDGDSAGEGGERA 180
D 121 LAQRAAHTFLIHGSRFYSSEARESNRARAFLALGMDGPDGDSAGEGGERA 180
QY 121 LAQRAAHTFLIHGSRFYSSEARESNRARAFLALGMDGPDGDSAGEGGERA 180
D 121 LAQRAAHTFLIHGSRFYSSEARESNRARAFLALGMDGPDGDSAGEGGERA 180
QY 181 APAGDAAAAGSGAFGAGGGAARCGGAAPLSPGATVALLLPAGPFLMLWGLAAGAR 240
D 181 APAGDAAAAGSGAFGAGGGAARCGGAAPLSPGATVALLLPAGPFLMLWGLAAGAR 240
QY 241 TAFVPTALRRGPIHCLRCGARAALVLAPEFLESLEPDLPALRAMGLHMAAGPGRHAG 300
D 241 TAFVPTALRRGPIHCLRCGARAALVLAPEFLESLEPDLPALRAMGLHMAAGPGRHAG 300
QY 301 ISDLAEVAEAVDGPVGYLSSPOSITDTCUYIFTSIGTGLPKAARISHKLILQCGFYQ 360
D 301 ISDLAEVAEAVDGPVGYLSSPOSITDTCUYIFTSIGTGLPKAARISHKLILQCGFYQ 360
QY 301 ISDLAEVAEAVDGPVGYLSSPOSITDTCUYIFTSIGTGLPKAARISHKLILQCGFYQ 360
D 301 ISDLAEVAEAVDGPVGYLSSPOSITDTCUYIFTSIGTGLPKAARISHKLILQCGFYQ 360
QY 361 LCGVHOEDVITYALPIYHWSGSLGVGCMGIGATVLLSKRSAGQFMDCCOHRVTVQ 420
D 361 LCGVHOEDVITYALPIYHWSGSLGVGCMGIGATVLLSKRSAGQFMDCCOHRVTVQ 420
QY 421 YIGELCRYLVNQPSPKABERGHKRLAVSGRLPDTWERFVRFGPLQVLETYGLTEGNA 480
D 421 YIGELCRYLVNQPSPKABERGHKRLAVSGRLPDTWERFVRFGPLQVLETYGLTEGNA 480
QY 481 TINYTGQRAVAGASWLYKHIFPFSILIRDYVTGEBIRPOQHCAWTSGEGLVAPVS 540
D 481 TINYTGQRAVAGASWLYKHIFPFSILIRDYVTGEBIRPOQHCAWTSGEGLVAPVS 540
QY 541 QOSPFLGYAGPBLAAGKLLKDVFRPGDVFPNTGDLVDCDDGFLRFHRTDITFRMKE 600
D 541 QOSPFLGYAGPBLAAGKLLKDVFRPGDVFPNTGDLVDCDDGFLRFHRTDITFRMKE 600
QY 601 NVATTEVAEVEALDFLQEVNNTYGVTPGHESRAGMAALVLRPHALDLMQLYTHSENL 660
D 601 NVATTEVAEVEALDFLQEVNNTYGVTPGHESRAGMAALVLRPHALDLMQLYTHSENL 660
QY 661 PPAARPRFLRLOESLATTETFKQOKRVMANEGPDSTLSDPYVLDQAVGAVLPLTARY 720
D 661 PPAARPRFLRLOESLATTETFKQOKRVMANEGPDSTLSDPYVLDQAVGAVLPLTARY 720
```

DB 661 PYYAPRRLROESLTTTETFKQKVRMANEGFDSTLSDPLVYLDQVAGVLPITTRY 720
 QY 721 SALLAGNLRI 730
 DB 721 SALLAGNLRI 730

RESULT 7
 ID ABU72207 standard; protein; 730 AA.
 XX
 AC ABU72207;
 XX
 DT 16-JUN-2003 (first entry)
 XX
 DE Novel human secreted and transmembrane protein PRO703.
 XX
 KW Human; secreted and transmembrane protein; PRO; antiinflammatory;
 KW antiarteriosclerotic; cardiatic; anti-infertility; anti-HIV; cytosclastic;
 KW antidiabetic; gene therapy; inflammatory disease; organ failure;
 KW atherosclerosis; cardiac injury; infertility; birth defect;
 KW premature aging; AIDS; cancer; diabetic complication; chromosome mapping;
 KW gene mapping; pharmaceutical; diagnostic; biosensor; bioreactor;
 KW tissue typing.
 OS Homo sapiens.
 XX
 PN US2002192706-A1.
 XX
 PD 19-DEC-2002.
 XX
 PF 24-OCT-2001; 2001US-00999832.
 XX
 PR 17-OCT-1997; 97US-0062250F.
 PR 13-NOV-1997; 97US-0064249F.
 PR 21-NOV-1997; 97US-0065311P.
 PR 10-MAR-1998; 98US-0077450P.
 PR 11-MAR-1998; 98US-0077632P.
 PR 11-MAR-1998; 98US-0077641P.
 PR 11-MAR-1998; 98US-0077649P.
 PR 13-MAR-1998; 98US-0077791P.
 PR 17-MAR-1998; 98US-0078004P.
 PR 20-MAR-1998; 98US-0078886P.
 PR 20-MAR-1998; 98US-0078910P.
 PR 20-MAR-1998; 98US-0078936P.
 PR 20-MAR-1998; 98US-0078939P.
 PR 25-MAR-1998; 98US-0079294P.
 PR 26-MAR-1998; 98US-0079656P.
 PR 27-MAR-1998; 98US-0079663P.
 PR 27-MAR-1998; 98US-0079664P.
 PR 27-MAR-1998; 98US-0079689P.
 PR 27-MAR-1998; 98US-0079728P.
 PR 27-MAR-1998; 98US-0079786P.
 PR 30-MAR-1998; 98US-0079920P.
 PR 30-MAR-1998; 98US-0079923P.
 PR 31-MAR-1998; 98US-0080105P.
 PR 31-MAR-1998; 98US-0080107P.
 PR 31-MAR-1998; 98US-0080165P.
 PR 31-MAR-1998; 98US-0080194P.
 PR 01-APR-1998; 98US-0080327P.
 PR 01-APR-1998; 98US-0080328P.
 PR 01-APR-1998; 98US-0080333P.
 PR 01-APR-1998; 98US-0080334P.
 PR 08-APR-1998; 98US-0081049P.
 PR 08-APR-1998; 98US-0081070P.
 PR 08-APR-1998; 98US-0081071P.
 PR 09-APR-1998; 98US-0081155P.
 PR 09-APR-1998; 98US-0081203P.
 PR 09-APR-1998; 98US-0081229P.
 PR 15-APR-1998; 98US-0081817P.

PR 15-APR-1998; 98US-0081819P.
 PR 15-APR-1998; 98US-0081838P.
 PR 15-APR-1998; 98US-0081952P.
 PR 15-APR-1998; 98US-0081955P.
 PR 21-APR-1998; 98US-0082568P.
 PR 21-APR-1998; 98US-0082569P.
 PR 22-APR-1998; 98US-0082700P.
 PR 22-APR-1998; 98US-0082704P.
 PR 22-APR-1998; 98US-0082797P.
 PR 22-APR-1998; 98US-0082804P.
 PR 23-APR-1998; 98US-0082796P.
 PR 07-OCT-1998; 98WO-US021141.
 PR 20-NOV-1998; 98WO-US024855.
 PR 08-MAR-1999; 99WO-US000106.
 PR 10-MAR-1999; 99WO-US000528.
 PR 14-MAY-1999; 99WO-US010733.
 PR 02-JUN-1999; 99WO-US012252.
 PR 30-NOV-1999; 99WO-US028313.
 PR 02-DEC-1999; 99WO-US028551.
 PR 02-DEC-1999; 99WO-US028565.
 PR 16-DEC-1999; 99WO-US030095.
 PR 30-DEC-1999; 99WO-US031243.
 PR 05-JAN-2000; 99WO-US031274.
 PR 06-JAN-2000; 2000WO-US000219.
 PR 06-JAN-2000; 2000WO-US000277.
 PR 11-FEB-2000; 2000WO-US000376.
 PR 18-FEB-2000; 2000WO-US003565.
 PR 24-FEB-2000; 2000WO-US004341.
 PR 02-MAR-2000; 2000WO-US005841.
 PR 10-MAR-2000; 2000WO-US006319.
 PR 21-MAR-2000; 2000WO-US007532.
 PR 30-MAR-2000; 2000WO-US008439.
 PR 17-MAY-2000; 2000WO-US013705.
 PR 22-MAY-2000; 2000WO-US014942.
 PR 30-MAY-2000; 2000WO-US014941.
 PR 02-JUN-2000; 2000WO-US015264.
 PR 28-JUL-2000; 2000WO-US020710.
 PR 24-AUG-2000; 2000WO-US023278.
 PR 01-DEC-2000; 2000WO-US032678.
 PR 20-DEC-2000; 2000WO-US034956.
 PR 28-FEB-2001; 2001WO-US006520.
 PR 22-MAR-2001; 2001WO-US009552.
 PR 25-MAY-2001; 2001WO-US017092.
 PR 01-JUN-2001; 2001WO-US017800.
 PR 20-JUN-2001; 2001WO-US019692.
 PR 29-JUN-2001; 2001WO-US021065.
 PR 09-JUL-2001; 2001WO-US021735.

XX
 PA (GENTH) GENENTECH INC.
 XX
 PI Ashkenazi AJ, Baker KP, Bolstein D, Desnoyers L, Eaton DL;
 PI Ferrara N, Filvaroff E, Fong S, Gao W, Geirder H, Gerritsen ME;
 PI Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan MJ;
 PI Kijavrin JJ, Kuo SS, Napier MA, Pan J, Paoni NF, Roy MA, Shelton DL;
 PI Stewart TA, Tumas D, Williams PM, Wood WI;
 XX
 DR WPI; 2003-328660/31.
 DR N-PSDB; ACA63545.
 XX
 PT New secreted and transmembrane nucleic acids and polypeptides, designated
 PT as PRO, useful for treating inflammation, organ failure, atherosclerosis,
 PT cardiac injury, infertility, birth defects, premature aging, AIDS, or
 PT cancer.
 PT
 PS Claim 12, Fig 39; 453pp; English.
 XX
 CC The invention describes an isolated nucleic acid (i) comprising, or which
 CC is at least 80 % sequence identity to, or the full-length coding sequence
 CC of, any of 118 300-2100 nucleotide sequences, which encodes its
 CC corresponding PRO polypeptide selected from 118 100-700 amino acid
 CC sequences, all given in the specification. The nucleic acids and

CC polypeptides are useful for treating inflammatory diseases, organ
 CC failure, atherosclerosis, cardiac injury, infertility, birth defects,
 CC premature aging, AIDS, cancer, or diabetic complications. The nucleic
 CC acids are useful as hybridisation probes, in chromosome and gene mapping,
 CC and in generating antisense RNA or DNA. The polypeptides are useful as
 CC pharmaceuticals, diagnostics, biosensors or bioreactors. Both are useful
 CC in tissue typing. This is the amino acid sequence of a novel human
 CC secreted and transmembrane PRO polypeptide
 CC
 XX

XX Sequence 730 AA:

Query Match 100.0%; Score 3843; DB 6; Length 730;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVCQRTAPAPWKEKSQLERAAIGFRKGGSGMFASGMNQTVPIDEGSMALLPLILL 60
 DB 1 MGVCQRTAPAPWKEKSQLERAAIGFRKGGSGMFASGMNQTVPIDEGSMALLPLILL 60
 QY 61 PLLIKLHLMPQLRWLPADIAFAVRALCCRRALRARAALAAAADPEGEGCSIAWRALAE 120
 DB 61 PLLIKLHLMPQLRWLPADIAFAVRALCCRRALRARAALAAAADPEGEGCSIAWRALAE 120
 QY 121 LAOQRAAHTFLIHGSRFRFSYSEARESNRARAFLRALGMDMGPDGDSGEGSAGEGERA 180
 DB 121 LAOQRAAHTFLIHGSRFRFSYSEARESNRARAFLRALGMDMGPDGDSGEGSAGEGERA 180
 QY 181 APGAGDAAGAGAEFAGDGAARGGAAAPISPGATVALLPAGPEFLIMFGLAKAGLR 240
 DB 181 APGAGDAAGAGAEFAGDGAARGGAAAPISPGATVALLPAGPEFLIMFGLAKAGLR 240
 QY 241 TAFVPTARRBPILHCHFCSCGARALYLAPEFLESLEPRLPRLMAGLHMAAGSTPAG 300
 DB 241 TAFVPTARRBPILHCHFCSCGARALYLAPEFLESLEPRLPRLMAGLHMAAGSTPAG 300
 QY 301 ISDLIAEVSADVGDPVPGYSSPISITDTCLYIFTSQTTGPKXARISHLILCOQGRYQ 360
 DB 301 ISDLIAEVSADVGDPVPGYSSPISITDTCLYIFTSQTTGPKXARISHLILCOQGRYQ 360
 QY 361 LCGHQBDEVTYLAFLPHMSGSLGIVCGMGIGATVVKSRFSAGQFVEDCQGRVYFQ 420
 DB 361 LCGHQBDEVTYLAFLPHMSGSLGIVCGMGIGATVVKSRFSAGQFVEDCQGRVYFQ 420
 QY 421 YIGELCRILVNOPPKAKRGHKVRLAVSGSLRPTDWERFVRFGLQVLETYGHTGNA 480
 DB 421 YIGELCRILVNOPPKAKRGHKVRLAVSGSLRPTDWERFVRFGLQVLETYGHTGNA 480
 QY 481 TINYTGQGAAGRASMLYKHIFPESLIYDYVTGEPIDPOGHCMATSPEGGLVAVS 540
 DB 481 TINYTGQGAAGRASMLYKHIFPESLIYDYVTGEPIDPOGHCMATSPEGGLVAVS 540
 QY 541 QOSPELGAAGPELAOGGLKDVFRPDPVFNPTGDLVCDQOGLRFRHDTGDTFRKKG 600
 DB 541 QOSPELGAAGPELAOGGLKDVFRPDPVFNPTGDLVCDQOGLRFRHDTGDTFRKKG 600
 QY 601 NVATTEVAEVEFALDFOEVNVYGVTVGHEGRAGMAALVLRPHALDLMQJYTHVSENL 660
 DB 601 NVATTEVAEVEFALDFOEVNVYGVTVGHEGRAGMAALVLRPHALDLMQJYTHVSENL 660
 QY 661 PPYAPRRLLOESLATTETPRQCKVRMANAGFDSITSDLYLDDQAVGAYLPLTARY 720
 DB 661 PPYAPRRLLOESLATTETPRQCKVRMANAGFDSITSDLYLDDQAVGAYLPLTARY 720
 QY 721 SALLAGNRI 730
 DB 721 SALLAGNRI 730
 AC ADA84115;
 AC ADA84115; standard; protein; 730 AA.

XX 20-NOV-2003 (first entry)
 XX Human POW124 protein.
 DE

KW human; marker; expressed sequence tag; EST; arabisopsis; tumour;
 KW stress-induced phenotype; hyperosmotic stress; colon cancer; immunogen;
 KW vaccine.
 XX

XX Homo sapiens.
 OS

PN NC2002103028-A2.
 XX

PD 27-DEC-2002.
 XX

PF 30-MAY-2002; 2002MO-1B004189.
 XX

PR 30-MAY-2001; 2001US-0293999P.
 XX

PR 22-OCT-2001; 2001US-0330457P.
 XX

PR 19-FEB-2002; 2002US-0357144P.
 XX

PA (BIOM-) BIO MEDICAL CENT.
 XX

PI Baranova AV, Yankovsky NK, Kozlov AP, Lobashev AV, Krukovskaya LL;
 DR WPI; 2003-175241/17.
 XX

DR N-PSDB; ADA84114.
 XX

PT Determining if a nucleic acid is a marker for a phenotype/cell type of
 PT interest, by global comparison of expressed sequence tags known to be
 PT expressed in the phenotype/cell type with all ESTs expressed in normal
 PT tissue.
 PS

PS Claim 29; Page 491-494; 516pp; English.
 XX

XX The invention relates to a novel method for determining if a nucleic acid
 CC is a marker for a predetermined phenotype/cell type of interest from a
 CC biological species. The method comprises performing a global comparison
 CC of a group of expressed sequence tags (ESTs) known to be expressed in the
 CC phenotype/cell type of interest with all ESTs expressed in normal tissue
 CC in order to identify ESTs that are preferentially expressed in the
 CC phenotype/cell type of interest. A method of the invention is useful for
 CC determining whether a nucleic acid is a marker for a predetermined
 CC phenotype or cell type of interest from a biological species, preferably
 CC Arabidopsis or human. The cell type of interest is an abnormal cell such
 CC as a tumour cell, and the predetermined phenotype is a stress-induced
 CC phenotype such as hyperosmotic stress or high salt conditions. A method
 CC of the invention is also useful for determining the progression of colon
 CC cancer in a human, for detecting a tumour cell, and for regulating or
 CC preventing the growth of a tumour cell. An antibody of the invention is
 CC useful for detecting the absence or presence of peptides encoded by
 CC tumour-associated markers. A polypeptide of the invention is useful as an
 CC immunogen for vaccinating an animal. The present sequence represents a
 CC tumour-associated antigen of the invention.
 XX

XX Sequence 730 AA:

Query Match 100.0%; Score 3843; DB 6; Length 730;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVCQRTAPAPWKEKSQLERAAIGFRKGGSGMFASGMNQTVPIDEGSMALLPLILL 60
 DB 1 MGVCQRTAPAPWKEKSQLERAAIGFRKGGSGMFASGMNQTVPIDEGSMALLPLILL 60
 QY 61 PLLIKLHLMPQLRWLPADIAFAVRALCCRRALRARAALAAAADPEGEGCSIAWRALAE 120
 DB 61 PLLIKLHLMPQLRWLPADIAFAVRALCCRRALRARAALAAAADPEGEGCSIAWRALAE 120
 QY 121 LAOQRAAHTFLIHGSRFRFSYSEARESNRARAFLRALGMDMGPDGDSGEGSAGEGERA 180
 DB 121 LAOQRAAHTFLIHGSRFRFSYSEARESNRARAFLRALGMDMGPDGDSGEGSAGEGERA 180

```
QY 181 ARGAGDAAGSGAEFGAGDGAARGGGAAPISPGATVALLLPAGEPFIMLWFLAKAGLR 240
DB 181 ARGAGDAAGSGAEFGAGDGAARGGGAAPISPGATVALLLPAGEPFIMLWFLAKAGLR 240
QY 241 TAFVPTALRRGPHLHCRSCGARALVLAPEFLESLEPDLPALRANGHLMAAGPQTDPAG 300
DB 241 TAFVPTALRRGPHLHCRSCGARALVLAPEFLESLEPDLPALRANGHLMAAGPQTDPAG 300
QY 301 ISDLLAEVSAEVDGPPGGLSSPDSITDTCLYIFTSQTGTPKARISHLXILCOQGFYQ 360
DB 301 ISDLLAEVSAEVDGPPGGLSSPDSITDTCLYIFTSQTGTPKARISHLXILCOQGFYQ 360
QY 361 LCGVHOEDVLYLALPLVHNSGSLIGVCMGIGATVVLKSFSAQFMEDCQHRVTFQ 420
DB 361 LCGVHOEDVLYLALPLVHNSGSLIGVCMGIGATVVLKSFSAQFMEDCQHRVTFQ 420
QY 421 YIGELCRVLYNOPPSKAERGHKVRILAVSGLRPDMERFVRFGLOVLETFGLTEGNA 480
DB 421 YIGELCRVLYNOPPSKAERGHKVRILAVSGLRPDMERFVRFGLOVLETFGLTEGNA 480
QY 481 TINYTGORGAVRASVLYGHIFFPSLIRYDTTGGPIRDPOGCMATSPGEPGLVAVS 540
DB 481 TINYTGORGAVRASVLYGHIFFPSLIRYDTTGGPIRDPOGCMATSPGEPGLVAVS 540
QY 541 QOSPFLGIVAGPEBLAQKLLKDVPRPGDVFFNTGDLVLCDDQGFLEPHDRGDTFRKMG 600
DB 541 QOSPFLGIVAGPEBLAQKLLKDVPRPGDVFFNTGDLVLCDDQGFLEPHDRGDTFRKMG 600
QY 601 NVATTEVAVFEPALDPLQEVNYYGTVPHGEGRAGMALVLRPPALMLQLYTHVSNTL 660
DB 601 NVATTEVAVFEPALDPLQEVNYYGTVPHGEGRAGMALVLRPPALMLQLYTHVSNTL 660
QY 661 PPYARRRFLRLQESLATTETFRQOKVRMANEGFDSSTLSDPLYLDQAVGYLPRTARY 720
DB 661 PPYARRRFLRLQESLATTETFRQOKVRMANEGFDSSTLSDPLYLDQAVGYLPRTARY 720
QY 721 SALLAGNRI 730
DB 721 SALLAGNRI 730

RESULT 9
ID ABU84887 standard; protein; 730 AA.
AC ABU84887;
XX
XX
DE 12-AUG-2003 (first entry)
XX
XX
KW Human secreted and transmembrane polypeptide PRO703.
XX
KW Human; thrombolytic agent; interferon; interleukin; cytokine;
KW erythropoietin; colony stimulating factor; cancer; colorectal carcinoma;
KW apoptosis related condition; AIDS; amyotrophic lateral sclerosis;
KW inflammatory disease; asthma; atherosclerosis; neurodegenerative disease;
KW gastrointestinal disorder; Alzheimer's disease; Parkinson's disease;
KW hypertension; myocardial ischemia; kidney disease; carcinoma;
KW glomerulonephritis; lung disease; pulmonary hypertension; preeclampsia;
KW bronchial asthma; gastric ulcer; renal failure; cardiovascular disease;
KW inflammatory bowel disease; reproductive disorder; premature labour.
XX
OS Homo sapiens.
XX
PN US2002177553-A1.
XX
PD 28-NOV-2002.
XX
PF 15-OCT-2001; 2001US-00978192.
XX
PR 17-OCT-1997; 97US-0062250P.
PR 03-NOV-1997; 97US-0064249P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0065364P.
```

PR 10-MAY-2001; 2001US-00854280.
 PR 25-MAY-2001; 2001WO-US017092.
 PR 01-JUN-2001; 2001US-00872035.
 PR 01-JUN-2001; 2001WO-US017800.
 PR 05-JUN-2001; 2001US-00874503.
 PR 14-JUN-2001; 2001US-00882636.
 PR 19-JUN-2001; 2001US-00886342.
 PR 20-JUN-2001; 2001WO-US019692.
 PR 29-JUN-2001; 2001WO-US021066.
 PR 09-JUL-2001; 2001WO-US021735.
 PR 30-JUL-2001; 2001US-00918585.
 (GETH) GENENTECH INC.
 XX
 PI Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL,
 PI Ferrara N, Filvarski PJ, Fong S, Gao W, Gerber H, Gertlisen ME,
 PI Goddard A, Godowski PJ, Grimaldi UC, Gunney AL, Hillan KJ,
 PI Klavin ID, Kuo SS, Napier MA, Pan J, Paoni NF, Roy MA, Shelton DL,
 PI Stewart TA, Tumas D, Williams PM, Wood WI;
 DR MPI; 2003-328499/31.
 DR N-PSDB; ACA71709.
 XX
 PT New isolated PRO polypeptides e.g. PRO213, PRO274 and PRO300, for use as
 PT pharmaceuticals, diagnostics, biosensors and bioreactors, for identifying
 PT modulators of receptor-ligand interactions.
 PS
 XX
 PS Claim 12; SEQ ID NO 102; 55bp; English.

CC The invention relates to an isolated secreted and transmembrane
 CC polypeptide, designated as PRO polypeptide. The PRO polypeptide is useful
 CC in PRO polypeptide detection methods. The PRO polypeptide is useful for
 CC linking a bioactive molecule to a cell. The PRO polypeptide or an
 CC antibody against it is useful for modulating a biological activity of a
 CC cell. The PRO polypeptide is useful in industrial applications including
 CC pharmaceuticals, diagnostics, biosensors and bioreactors. The PRO
 CC polypeptide is also useful as a thrombolytic agent, interferon,
 CC interleukin, erythropoietin, colony stimulating factor and other
 CC cytokines. The PRO polypeptide is useful for treating disease such as
 CC cancer e.g. colorectal carcinoma; apoptosis related conditions e.g. AIDS,
 CC atherosclerotic lateral sclerosis; inflammatory disease e.g. asthma,
 CC atherosclerosis; neurodegenerative disease e.g. Alzheimer's disease,
 CC Parkinson's disease; cardiovascular disease e.g. hypertension and
 CC myocardial ischaemia; kidney disease e.g. renal failure and
 CC glomerulonephritis; lung disease e.g. pulmonary hypertension, bronchial
 CC asthma; gastrointestinal disorders e.g. gastric ulcer and inflammatory
 CC bowel disease; reproductive disorders e.g. premature labour and
 CC preclampsia; carcinogenesis. The present sequence represents the amino
 CC acid sequence of a PRO polypeptide of the invention. Note: The sequence
 CC data for this patent did not form part of the printed specification but
 CC was obtained in electronic format directly from USPTO at
 CC segdate.uspto.gov/sequence.html?docID=20020177553
 CC
 XX
 XX Sequence 730 AA;

Query Match 100.0%; Score 3843; DB 6; Length 730;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVCORTAPMKESQLEKRAALGFRKGGSGMFASGMQTVIEEAGMAALLPLLLLL 60
 DB 1 MGVCORTAPMKESQLEKRAALGFRKGGSGMFASGMQTVIEEAGMAALLPLLLLL 60
 QY 61 PLLILKILMPLRPLADLAFAPVALCKKALPRAALAAAAPBEGGCSLAMPRLAE 120
 DB 61 PLLILKILMPLRPLADLAFAPVALCKKALPRAALAAAAPBEGGCSLAMPRLAE 120
 QY 121 LAQGRRAHTFLHSGRRFSYSEARESNRAARAFIRALGMDWPGDGSAGSGEGERA 180
 DB 121 LAQGRRAHTFLHSGRRFSYSEARESNRAARAFIRALGMDWPGDGSAGSGEGERA 180
 QY 181 AFGAGDAAAGSAGAFAGDGAARGGGAAAPLSGATVALLPAGPEFLMLFGLAKGLR 240
 DB 181 AFGAGDAAAGSAGAFAGDGAARGGGAAAPLSGATVALLPAGPEFLMLFGLAKGLR 240

DB 181 AFGAGDAAAGSAGAFAGDGAARGGGAAAPLSGATVALLPAGPEFLMLFGLAKGLR 240
 QY 241 TAFVPTALRRGPIHLHLRSCGARALYLAPPELSLEPDLPAIRAMGLHMAAGPCHPAG 300
 DB 241 TAFVPTALRRGPIHLHLRSCGARALYLAPPELSLEPDLPAIRAMGLHMAAGPCHPAG 300
 QY 301 ISDLAEVSAEAVGAPVPGYLSPOSITPDLCLVITFTSGTGLPAAARISHLKILQCGGFYQ 360
 DB 301 ISDLAEVSAEAVGAPVPGYLSPOSITPDLCLVITFTSGTGLPAAARISHLKILQCGGFYQ 360
 QY 361 LCGVHEDVIYIALPLPHNNGSLIGVCMGIGATVILSKFSAGQFWDCCQHRVTYFQ 420
 DB 361 LCGVHEDVIYIALPLPHNNGSLIGVCMGIGATVILSKFSAGQFWDCCQHRVTYFQ 420
 QY 421 YIGELGRVYNOPPSAERGHVYRLAVSGSLPDMFERVRRFPGVOVETGLTEGNVA 480
 DB 421 YIGELGRVYNOPPSAERGHVYRLAVSGSLPDMFERVRRFPGVOVETGLTEGNVA 480
 QY 481 TINYTGORGAVGRASWLYKHIFPFSILRYDVTTGEBIRDPOGHMAATSPDEGLVAPVS 540
 DB 481 TINYTGORGAVGRASWLYKHIFPFSILRYDVTTGEBIRDPOGHMAATSPDEGLVAPVS 540
 QY 541 QOSPFGVAGGPELLAAGKLLKDYRRGDPVFPMTGDLVNDGGLRPHRTGDPFKWKE 600
 DB 541 QOSPFGVAGGPELLAAGKLLKDYRRGDPVFPMTGDLVNDGGLRPHRTGDPFKWKE 600
 QY 601 NVATTEVAEVEFALDLQEVNVYGVTPGHESRAGMAALVLRPHALDLMQLYTHVSENL 660
 DB 601 NVATTEVAEVEFALDLQEVNVYGVTPGHESRAGMAALVLRPHALDLMQLYTHVSENL 660
 QY 661 PYPARPPRLRLOSLATTTTFCQKQRYANEGEDBSTLSDPLVLDQANGAHLPTTAY 720
 DB 661 PYPARPPRLRLOSLATTTTFCQKQRYANEGEDBSTLSDPLVLDQANGAHLPTTAY 720
 QY 721 SALLAGNLRI 730
 DB 721 SALLAGNLRI 730
 RESULT 10
 ID ABU61085 standard; protein; 730 AA.
 XX ABU61085;
 XX 08-MAY-2003 (first entry)
 XX
 DE Human PRO703 polypeptide.
 XX
 KM Human; PRO polypeptide; secreted and transmembrane protein;
 KM immune disorder; diabetes; hyper-insulinaemia; hypo-insulinaemia;
 KM cardiac insufficiency; nervous system disorder; kidney disorder;
 KM bone disorder; cartilage disorder; arthritis; tumour; wound healing;
 KM genetic disorder; cytostatic; antidiabetic; antiinflammatory;
 KM antiarthritic; anti-tumour; vulnery; antianaemic; dermatological;
 KM cardiant.
 XX
 XX Homo sapiens.
 XX OS
 XX US2002169284-A1.
 XX PD 14-NOV-2002.
 XX
 PF 16-OCT-2001; 2001US-00978697.
 XX
 XX 26-MAY-1981; 81US-00267213.
 XX 17-OCT-1997; 97US-0062250P.
 XX 03-NOV-1997; 97US-0064249P.
 XX 13-NOV-1997; 97US-0065311P.
 XX 21-NOV-1997; 97US-0066364P.
 XX 10-MAR-1998; 98US-0077450P.
 XX 11-MAR-1998; 98US-0077632P.
 XX 11-MAR-1998; 98US-0077641P.

PR 11-MAR-1998; 98US-007649P.
 PR 12-MAR-1998; 98US-0077791P.
 PR 13-MAR-1998; 98US-0078004P.
 PR 17-MAR-1998; 98US-00040220.
 PR 20-MAR-1998; 98US-0078886P.
 PR 20-MAR-1998; 98US-0078910P.
 PR 20-MAR-1998; 98US-0078936P.
 PR 20-MAR-1998; 98US-0078939P.
 PR 25-MAR-1998; 98US-0079254P.
 PR 26-MAR-1998; 98US-0079656P.
 PR 27-MAR-1998; 98US-0079663P.
 PR 27-MAR-1998; 98US-0079664P.
 PR 27-MAR-1998; 98US-0079689P.
 PR 27-MAR-1998; 98US-0079728P.
 PR 30-MAR-1998; 98US-0079786P.
 PR 30-MAR-1998; 98US-0079920P.
 PR 26-JUN-1998; 98US-00105413.
 PR 07-OCT-1998; 98US-0016897P.
 PR 07-OCT-1998; 98US-00202114.
 PR 02-NOV-1998; 98US-00184216.
 PR 06-NOV-1998; 98US-00187368.
 PR 20-NOV-1998; 98US-00204855.
 PR 07-DEC-1998; 98US-00202054.
 PR 22-DEC-1998; 98US-00218517.
 PR 05-JAN-1999; 98US-00000106.
 PR 05-MAR-1999; 98US-00254465.
 PR 08-MAR-1999; 98US-00050502.
 PR 10-MAR-1999; 98US-00265686.
 PR 10-MAR-1999; 98US-00050519.
 PR 12-APR-1999; 98US-00284291.
 PR 14-MAY-1999; 98US-00311832.
 PR 14-MAY-1999; 98US-00310733.
 PR 02-JUN-1999; 98US-00312252.
 PR 25-AUG-1999; 98US-00380137.
 PR 25-AUG-1999; 98US-00380138.
 PR 25-AUG-1999; 98US-00380138.
 PR 30-NOV-1999; 98US-00380142.
 PR 02-DEC-1999; 98US-00283313.
 PR 02-DEC-1999; 98US-00285551.
 PR 02-DEC-1999; 98US-00285555.
 PR 16-DEC-1999; 98US-00300055.
 PR 30-DEC-1999; 98US-0031243.
 PR 30-DEC-1999; 98US-0031274.
 PR 05-JAN-2000; 2000US-0000219.
 PR 06-JAN-2000; 2000US-0000227.
 PR 11-FEB-2000; 2000US-0000376.
 PR 11-FEB-2000; 2000US-00003565.
 PR 18-FEB-2000; 2000US-0004341.
 PR 24-FEB-2000; 2000US-0005004.
 PR 02-MAR-2000; 2000US-0005841.
 PR 10-MAR-2000; 2000US-0006319.
 PR 21-MAR-2000; 2000US-0007532.
 PR 30-MAR-2000; 2000US-0008439.
 PR 17-MAY-2000; 2000US-0013705.
 PR 22-MAY-2000; 2000US-0014042.
 PR 30-JUN-2000; 2000US-0014941.
 PR 02-JUN-2000; 2000US-0015264.
 PR 28-JUL-2000; 2000US-0020710.
 PR 24-AUG-2000; 2000US-0023328.
 PR 08-NOV-2000; 2000US-00709238.
 PR 27-NOV-2000; 2000US-00723749.
 PR 01-DEC-2000; 2000US-0032678.
 PR 20-DEC-2000; 2000US-00747253.
 PR 20-DEC-2000; 2000US-00747253.
 PR 28-FEB-2001; 2001US-0006520.
 PR 22-MAR-2001; 2001US-00816744.
 PR 22-MAR-2001; 2001US-00816920.
 PR 22-MAR-2001; 2001US-00816920.
 PR 10-MAY-2001; 2001US-00854208.
 PR 10-MAY-2001; 2001US-00854289.
 PR 25-MAY-2001; 2001US-00854289.
 PR 01-JUN-2001; 2001US-00872035.
 PR 01-JUN-2001; 2001US-00872035.

PR 05-JUN-2001; 2001US-00874503.
 PR 14-JUN-2001; 2001US-00882636.
 PR 19-JUN-2001; 2001US-00886342.
 PR 20-JUN-2001; 2001US-0089692.
 PR 23-JUN-2001; 2001US-00921066.
 PR 09-JUL-2001; 2001US-00921735.
 PR 30-JUL-2001; 2001US-00918585.
 XX
 XX (GENTH) GENENTECH INC.
 PA
 XX Ashkenazi A, Baker KP, Botstein D, Desnoyers J, Eaton D;
 PI Ferrara N, Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME;
 PI Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ;
 PI Kijavrin IJ, Kuo SS, Napier MA, Pan J, Paoletti NF, Roy MA, Shelton DL;
 PI Stewart TA, Tumas D, Williams PM, Wood WT;
 XX
 XX WPI; 2003-288163/28.
 DR N-PSDB; ABX92349.
 XX
 PT Novel secreted and transmembrane polypeptides and polynucleotides
 PT encoding them useful for treating cancer, kidney diseases, bone,
 PT cartilage disorders and immune deficiencies.
 XX
 XX Claim 12; Fig 39; 459pp; English.
 XX
 CC The present invention relates to the isolation of novel human PRO
 CC polypeptides, and the polynucleotide sequences encoding them. The PRO
 CC polypeptides are secreted and transmembrane proteins. The PRO
 CC polypeptides are useful for detecting other PRO polypeptides, for linking
 CC bioactive molecules to cells expressing PRO polypeptides, for modulating
 CC biological activities of cells expressing PRO polypeptides, and for for
 CC identifying agonists or antagonists. The bioactive molecule may be a
 CC toxin, radiolabel or antibody, and causes apoptosis or death of the cell.
 CC The PRO polypeptides are useful for treating immune disorders, diabetes
 CC or hyper- or hypo-insulinaemia, cardiac insufficiency, nervous system
 CC disorders, kidney disorders, bone and cartilage disorders or arthritis,
 CC tumours, and wound healing. The polynucleotide sequences encoding PRO
 CC polypeptides are useful as hybridisation probes, in chromosome and gene
 CC mapping, in the generation of antisense RNA and DNA, in the preparation
 CC of PRO polypeptides, for generating transgenic animals or knockout
 CC animals, for the genetic analysis of individuals with genetic disorders,
 CC and in gene therapy. ABUS1071-ABUS1616 represent the human PRO
 CC polypeptides of the invention. Note: The sequence data for this patent
 CC was obtained in electronic format directly from the USPTO web site at
 CC seqdata.uspto.gov/pspsdidentry.html
 XX
 XX Sequence 730 AA;
 SO
 Query Match 100.0%; Score 3843; DB 6; Length 730;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGVCORTRAPWKEKSLERAPALGFRKSGSGMFASGWNQTVPIEAGSMALLLPLLLLL 60
 Db 1 MGVCORTRAPWKEKSLERAPALGFRKSGSGMFASGWNQTVPIEAGSMALLLPLLLLL 60
 QY 61 PLLILKLIHMPOLRMLPADLPAVRLCCKRLRPARALAAAADDEGGGSLAWRLAE 120
 Db 61 PLLILKLIHMPOLRMLPADLPAVRLCCKRLRPARALAAAADDEGGGSLAWRLAE 120
 QY 121 LAQGRRAHFTLHNGSRFSYSERESNRRAAPAFALGMDGPPGCGSGEGSAGEGERRA 180
 Db 121 LAQGRRAHFTLHNGSRFSYSERESNRRAAPAFALGMDGPPGCGSGEGSAGEGERRA 180
 QY 121 LAQGRRAHFTLHNGSRFSYSERESNRRAAPAFALGMDGPPGCGSGEGSAGEGERRA 180
 Db 121 LAQGRRAHFTLHNGSRFSYSERESNRRAAPAFALGMDGPPGCGSGEGSAGEGERRA 180
 QY 181 APGAGDAAAGSGAEFAGDGAARGGGAAAPLSPGATVALLDPAGEPFLWFGAKAGR 240
 Db 181 APGAGDAAAGSGAEFAGDGAARGGGAAAPLSPGATVALLDPAGEPFLWFGAKAGR 240
 QY 241 TAFVPTALRGRLHGLRSCGARALVIAPEFLSLEPDI PARANGLHWAGPGTHPAG 300
 Db 241 TAFVPTALRGRLHGLRSCGARALVIAPEFLSLEPDI PARANGLHWAGPGTHPAG 300
 QY 301 ISDLIAEVSAYVDGVPVGYLSSPOSITDTCLYIFTSGLTGLPKAARISHKLIQCGGFYQ 360
 Db 301 ISDLIAEVSAYVDGVPVGYLSSPOSITDTCLYIFTSGLTGLPKAARISHKLIQCGGFYQ 360

30-JUL-2001; 2001US-00918565.

(GENTH) GENENTECH INC.

Ashkenazi AJ, Baker KP, Botstein D, Desnovers L, Baton DL, Ferrara N, Flvaroff E, Fong S, Gao W, Gerber H, Gertissen ME, Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ, Kijavini JY, Kuo SS, Napier MA, Par J, Pooni NF, Roy MA, Shelton DL, Stewart TA, Tumes D, Williams PM, Wood WJ, WPI, 2003-341189/32.

N-PFDB; ACA66090.

New genes and secreted and transmembrane polypeptides (e.g. PRO337 or PRO1559), useful for treating or diagnosing e.g. cancers, atherosclerosis, infertility, stroke, encephalitis, hepatitis or multiple sclerosis in mammals.

Claim 12; Fig 39; 460P; English.

The invention relates to a new isolated nucleic acid molecule comprising a sequence with at least 80% identity to: (a) a nucleotide encoding any of 94 PRO polypeptides whose sequences are fully defined in the specification; or (b) any of 94 nucleotide sequences fully defined in the specification; or the full length coding sequence of any these 94 nucleotide sequences. Also included are an isolated PRO polypeptide scoring at least 80% positives when compared to any of the PRO polypeptide sequences cited above (or an isolated PRO polypeptide having at least 80% amino acid sequence identity to: (a) an amino acid sequence encoded by the nucleotide deposited with ATCC numbers listed in the specification; (b) the PRO polypeptide, lacking its associated signal peptide; or (c) an extracellular domain of the PRO polypeptide, with or lacking its associated signal peptide), a vector comprising the nucleic acid molecule, a host cell comprising the vector (and producing a PRO polypeptide), a chimeric molecule comprising the PRO polypeptide fused to a heterologous amino acid sequence and an anti-PRO antibody. The PRO polypeptides or polynucleotides are useful as pharmaceuticals, diagnostics, biosensors or bioreactors. These are particularly useful for detecting or treating e.g. malignancies or cancers (e.g. ovarian cancer, colorectal cancer, sarcoma, leukaemia or lymphoma), inflammatory disease, necrosis, atherosclerosis, infertility, premature aging, psoriasis, inflammatory disease, renal disease, arthritis, immune-mediated alopecia, stroke, encephalitis, hepatitis, or multiple sclerosis in mammals. The PRO polypeptides are useful in drug screening, particularly as targets for therapeutic intervention in these diseases, and in the diagnostic determination of the presence of these diseases. The PRO polypeptides are also useful as molecular weight markers, or for chromosome identification. The PRO genes are useful as hybridisation probes, or for screening libraries of human cDNA, genomic DNA or mRNA. The PRO genes may also be used in gene therapy, particularly for replacing a defective gene. The present sequence represents a PRO polypeptide

Sequence 730 AA;

Query Match 100.0%; Score 3843; DB 6; Length 730;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 MGVCORTAPWKEKQLEBAALGFRKGGSGMFGSMNQTVPHEEGSMAALLPILLL 60
 1 MGVCORTAPWKEKQLEBAALGFRKGGSGMFGSMNQTVPHEEGSMAALLPILLL 60
 DB 61 PILLIKLHMPOLRLPADLFAVVALCCKALRRARALAAAADPEGEGCSLAWRLAE 120
 61 PILLIKLHMPOLRLPADLFAVVALCCKALRRARALAAAADPEGEGCSLAWRLAE 120
 DB 121 LAOQRAAHTFLIHGRRRSYSFAERESNRARAFTRALGMDWGPDDGSGSGSAGEGERA 180
 121 LAOQRAAHTFLIHGRRRSYSFAERESNRARAFTRALGMDWGPDDGSGSGSAGEGERA 180
 DB 121 LAOQRAAHTFLIHGRRRSYSFAERESNRARAFTRALGMDWGPDDGSGSGSAGEGERA 180
 121 LAOQRAAHTFLIHGRRRSYSFAERESNRARAFTRALGMDWGPDDGSGSGSAGEGERA 180
 DB 181 AFGADPAAAGSAEAFGAGDGAARCGAALPSPGATVALLPAGBEFLMIFGAKAGLR 240
 181 AFGADPAAAGSAEAFGAGDGAARCGAALPSPGATVALLPAGBEFLMIFGAKAGLR 240
 DB 181 AFGADPAAAGSAEAFGAGDGAARCGAALPSPGATVALLPAGBEFLMIFGAKAGLR 240

241 TAFVPTALRGPHLHCLRSRGARALVLAPEFLESLEPDLPALRAMGLHMAAGPTHPAG 300
 241 TAFVPTALRGPHLHCLRSRGARALVLAPEFLESLEPDLPALRAMGLHMAAGPTHPAG 300
 DB 301 ISDLAEVSAEVDGPPVPGILSSPQSTITDCLVITFTSGITGLPKAARISHUKLQCGFVQ 360
 301 ISDLAEVSAEVDGPPVPGILSSPQSTITDCLVITFTSGITGLPKAARISHUKLQCGFVQ 360
 DB 361 LCGVHOEDVITVLAPEYHMSGSLIGVCGMGICATVLSKKSAGOFMEDCOHRTVYQ 420
 361 LCGVHOEDVITVLAPEYHMSGSLIGVCGMGICATVLSKKSAGOFMEDCOHRTVYQ 420
 DB 421 YIGELCRYLVNPPSKAERGHKVRALVSGSLPDTWERFRRFGPLQVLETYGLTEGNA 480
 421 YIGELCRYLVNPPSKAERGHKVRALVSGSLPDTWERFRRFGPLQVLETYGLTEGNA 480
 DB 481 TINYTGQRAGVGRASLKYHTIPFSLIRVDYTTGETIRPQCHMATSPGEGLLVAPVS 540
 481 TINYTGQRAGVGRASLKYHTIPFSLIRVDYTTGETIRPQCHMATSPGEGLLVAPVS 540
 DB 541 QQSPFLGYAGPELQAGKLLKDFRPGDFVFTGDLVCDGQFLRFHRTGDFRMKGE 600
 541 QQSPFLGYAGPELQAGKLLKDFRPGDFVFTGDLVCDGQFLRFHRTGDFRMKGE 600
 DB 601 NVATTVEAVFEALDPLQEVNNTYGVTVPGHEGRAGMAALVLRPPALDMQLYTVSENL 660
 601 NVATTVEAVFEALDPLQEVNNTYGVTVPGHEGRAGMAALVLRPPALDMQLYTVSENL 660
 DB 661 PPVARRPFLQBSLATSTTFKQOKVRMANEGFDPSTLSDPYYVLDQAVGYLPTTARY 720
 661 PPVARRPFLQBSLATSTTFKQOKVRMANEGFDPSTLSDPYYVLDQAVGYLPTTARY 720
 DB 721 SALLAGNLRI 730
 721 SALLAGNLRI 730

RESULT 12
 ADA24641
 ID ADA24641 standard; protein; 730 AA.
 XX
 ADA24641;
 AC
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Novel human secreted and transmembrane protein PRO703.
 XX
 KM Human; secreted and transmembrane protein; PRO; tissue typing;
 KM chromosome identification; vaccine; cancer; retinal disorder;
 KM sports-related joint disorder; osteoarthritis; rheumatoid arthritis;
 KM wound healing; obesity; diabetes; hearing loss;
 KM cardiac insufficiency disorder; kidney disorder; nervous system disorder;
 KM haemoglobin associated disorder.
 KM
 OS Homo sapiens.
 PN US2003050241-A1.
 XX
 PD 13-MAR-2003.
 XX
 PF 16-OCT-2001; 2001US-00978564.
 XX
 PR 17-OCT-1997; 97US-0062250P.
 PR 03-NOV-1997; 97US-0064249P.
 PR 13-NOV-1997; 97US-0065311P.
 PR 21-NOV-1997; 97US-0066364P.
 PR 10-MAR-1998; 98US-0077450P.
 PR 11-MAR-1998; 98US-0077632P.
 PR 11-MAR-1998; 98US-0077641P.
 PR 11-MAR-1998; 98US-0077649P.
 PR 12-MAR-1998; 98US-0077791P.
 PR 13-MAR-1998; 98US-0078004P.

PR 20-MAR-1998; 98US-0078866P.
 PR 20-MAR-1998; 98US-0078910P.
 PR 20-MAR-1998; 98US-0078936P.
 PR 20-MAR-1998; 98US-0078939P.
 PR 25-MAR-1998; 98US-0079294P.
 PR 26-MAR-1998; 98US-0079656P.
 PR 27-MAR-1998; 98US-0079663P.
 PR 27-MAR-1998; 98US-0079666P.
 PR 27-MAR-1998; 98US-0079688P.
 PR 27-MAR-1998; 98US-0079728P.
 PR 27-MAR-1998; 98US-0079786P.
 PR 30-MAR-1998; 98US-0079920P.
 PR 30-MAR-1998; 98US-0079923P.
 PR 31-MAR-1998; 98US-0080105P.
 PR 31-MAR-1998; 98US-0080107P.
 PR 31-MAR-1998; 98US-0080165P.
 PR 31-MAR-1998; 98US-0080194P.
 PR 01-APR-1998; 98US-0080327P.
 PR 01-APR-1998; 98US-0080328P.
 PR 01-APR-1998; 98US-0080333P.
 PR 01-APR-1998; 98US-0080334P.
 PR 08-APR-1998; 98US-0081049P.
 PR 08-APR-1998; 98US-0081070P.
 PR 08-APR-1998; 98US-0081071P.
 PR 09-APR-1998; 98US-0081195P.
 PR 09-APR-1998; 98US-0081203P.
 PR 09-APR-1998; 98US-0081229P.
 PR 15-APR-1998; 98US-0081817P.
 PR 15-APR-1998; 98US-0081819P.
 PR 15-APR-1998; 98US-0081838P.
 PR 15-APR-1998; 98US-0081952P.
 PR 21-APR-1998; 98US-0082568P.
 PR 21-APR-1998; 98US-0082569P.
 PR 22-APR-1998; 98US-0082700P.
 PR 22-APR-1998; 98US-0082704P.
 PR 22-APR-1998; 98US-0082797P.
 PR 22-APR-1998; 98US-0082804P.
 PR 23-APR-1998; 98US-0082799P.
 PR 27-APR-1998; 98US-0083336P.
 PR 28-APR-1998; 98US-0083322P.
 PR 29-APR-1998; 98US-0083392P.
 PR 29-APR-1998; 98US-0083455P.
 PR 29-APR-1998; 98US-0083496P.
 PR 29-APR-1998; 98US-0083499P.
 PR 29-APR-1998; 98US-0083500P.
 PR 29-APR-1998; 98US-0083545P.
 PR 29-APR-1998; 98US-0083554P.
 PR 29-APR-1998; 98US-0083558P.
 PR 29-APR-1998; 98US-0083559P.
 PR 30-APR-1998; 98US-0083742P.
 PR 05-MAY-1998; 98US-0084366P.
 PR 06-MAY-1998; 98US-0084414P.
 PR 06-MAY-1998; 98US-0084441P.
 PR 07-MAY-1998; 98US-0084598P.
 PR 07-MAY-1998; 98US-0084600P.
 PR 07-MAY-1998; 98US-0084627P.
 PR 07-MAY-1998; 98US-0084637P.
 PR 07-MAY-1998; 98US-0084639P.
 PR 07-MAY-1998; 98US-0084640P.
 PR 07-MAY-1998; 98US-0084643P.
 PR 13-MAY-1998; 98US-0085323P.
 PR 13-MAY-1998; 98US-0085338P.
 PR 13-MAY-1998; 98US-0085339P.
 PR 15-MAY-1998; 98US-0085573P.
 PR 15-MAY-1998; 98US-0085579P.
 PR 15-MAY-1998; 98US-0085580P.
 PR 15-MAY-1998; 98US-0085582P.
 PR 15-MAY-1998; 98US-0085689P.
 PR 15-MAY-1998; 98US-0085697P.
 PR 15-MAY-1998; 98US-0085700P.
 PR 15-MAY-1998; 98US-0085704P.
 PR 18-MAY-1998; 98US-0086023P.

PR 22-MAY-1998; 98US-0086392P.
 PR 22-MAY-1998; 98US-0086414P.
 PR 22-MAY-1998; 98US-0086430P.
 PR 22-MAY-1998; 98US-0086486P.
 PR 28-MAY-1998; 98US-0087088P.
 PR 28-MAY-1998; 98US-0087106P.
 PR 28-MAY-1998; 98US-0087208P.
 PR 28-MAY-1998; 98US-0090863P.
 PR 25-JUN-1998; 98US-0091010P.
 PR 25-JUN-1998; 98US-0091012P.
 PR 01-JUL-1998; 98US-0091359P.
 PR 30-JUL-1998; 98US-0094651P.
 PR 11-SEP-1998; 98US-0100038P.
 PR 07-OCT-1998; 98MO-US021141.
 PR 20-NOV-1998; 98MO-US024855.
 PR 22-DEC-1998; 98US-0113296P.
 PR 22-DEC-1998; 98US-0113621P.
 PR 05-JAN-1999; 99MO-US000106.
 PR 08-MAR-1999; 99MO-US005028.
 PR 10-MAR-1999; 99MO-US005190.
 PR 12-MAR-1999; 99US-0123957P.
 PR 29-MAR-1999; 99US-0126773P.
 PR 21-APR-1999; 99US-0130232P.
 PR 26-APR-1999; 99US-0131022P.
 PR 28-APR-1999; 99US-0131445P.
 PR 14-MAY-1999; 99US-0134287P.
 PR 14-MAY-1999; 99MO-US010733.
 PR 02-JUN-1999; 99MO-US011252.
 PR 16-JUN-1999; 99US-0141037P.
 PR 23-JUN-1999; 99US-0145698P.
 PR 26-JUL-1999; 99US-0145698P.
 PR 28-JUL-1999; 99US-0146222P.
 PR 29-OCT-1999; 99US-0162506P.
 PR 30-NOV-1999; 99MO-US028313.
 PR 02-DEC-1999; 99MO-US028551.
 PR 02-DEC-1999; 99MO-US028565.
 PR 16-DEC-1999; 99MO-US030095.
 PR 30-DEC-1999; 99MO-US031243.
 PR 30-DEC-1999; 99MO-US031274.
 PR 05-JAN-2000; 2000MO-US000219.
 PR 06-JAN-2000; 2000MO-US000277.
 PR 06-JAN-2000; 2000MO-US000376.
 PR 11-FEB-2000; 2000MO-US003565.
 PR 18-FEB-2000; 2000MO-US004341.
 PR 24-FEB-2000; 2000MO-US005004.
 PR 02-MAR-2000; 2000MO-US005841.
 PR 10-MAR-2000; 2000MO-US006319.
 PR 21-MAR-2000; 2000MO-US007532.
 PR 30-MAR-2000; 2000MO-US008439.
 PR 17-MAY-2000; 2000MO-US013705.
 PR 22-MAY-2000; 2000MO-US014042.
 PR 30-MAY-2000; 2000MO-US014941.
 PR 02-JUN-2000; 2000MO-US015264.
 PR 28-JUL-2000; 2000MO-US020710.
 PR 24-AUG-2000; 2000MO-US023328.
 PR 01-DEC-2000; 2000MO-US032678.
 PR 20-DEC-2000; 2000MO-US034956.
 PR 28-FEB-2001; 2001MO-US006520.
 PR 22-MAR-2001; 2001MO-US009552.
 PR 25-MAY-2001; 2001MO-US017099.
 PR 01-JUN-2001; 2001MO-US017800.
 PR 20-JUN-2001; 2001MO-US019692.
 PR 29-JUN-2001; 2001MO-US021066.
 PR 09-JUL-2001; 2001MO-US021735.
 PR 30-JUL-2001; 2001US-00918585.

(GETH) GENENTECH INC.

XX Ashkenazi AJ, Baker KP, Bolstein D, Desnoyers L, Eaton DL;
 PI Bertara N, Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME;
 PI Goddard A, Godowski FO, Grimaldi JC, Gurney AL, Hillian KJ, Shelton DL;
 PI Kijavlin IJ, Kuo SS, Napier MA, Pan J, Paoni NF, Roy MA, Shelton DL;

PI Stewart TA, Tumas D, Williams PM, Wood WI;
XX WPI; 2003-521814/49.
DR N-PSDB; ADA24640.
XX
PT New isolated PRO polypeptides for example extracellular, secreted and
PT membrane bound proteins, useful for modulating the biological activities
PT of cells and for treating, for example diabetes, cancer, rheumatoid
PT arthritis, and hearing loss.
XX
XX
PS Claim 12; Fig 39; 461pp; English.

CC The invention describes an isolated secreted and transmembrane (PRO)
CC polypeptide (1). PRO337 polypeptide is useful for detecting PRO4993
CC polypeptide in a sample, and vice versa. PRO725, PRO700 and PRO739 are
CC useful for detecting PRO1559 polypeptide in a sample, and PRO1559 is
CC useful for detecting PRO725, PRO700 and PRO739 in a sample. PRO4993 is
CC useful for linking a bioactive molecule to a cell expressing a PRO337
CC polypeptide, and PRO337 is useful for linking a bioactive molecule to a
CC cell expressing a PRO4993 polypeptide. PRO1559 is useful for linking a
CC bioactive molecule to a cell expressing a PRO735, PRO700 and PRO739

Query Match 100.0%; Score 3843; DB 6; Length 730;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVCORTRAPMKEKSOLEPAALGPRKSGGMFASGNGTVPTEAGSMAALLPILLL 60
DB 1 MGVCORTRAPMKEKSOLEPAALGPRKSGGMFASGNGTVPTEAGSMAALLPILLL 60
QY 61 PILLKLMPLQRLMPLADLAFVAVLCCCKRLRRALAAAADPEGGGCSLAWRLAE 120
DB 61 PILLKLMPLQRLMPLADLAFVAVLCCCKRLRRALAAAADPEGGGCSLAWRLAE 120
QY 121 LMOQRAHTFLHSGRRSYSEARESNRAARAFRLGMDGPGGSGSGSGEGERA 180
DB 121 LMOQRAHTFLHSGRRSYSEARESNRAARAFRLGMDGPGGSGSGSGEGERA 180
QY 181 AEGAGDAAAGSAGEFAGDGAARCGAALPLSPGATVALLPAGPEFLMWFLAKAGLR 240
DB 181 AEGAGDAAAGSAGEFAGDGAARCGAALPLSPGATVALLPAGPEFLMWFLAKAGLR 240
QY 241 TAFVPAALRGPLILHSCGCAALVLAPEPLESEPLPALRANGMLMAAGGTTHAG 300
DB 241 TAFVPAALRGPLILHSCGCAALVLAPEPLESEPLPALRANGMLMAAGGTTHAG 300
QY 301 ISDLAEVSAEVDGVPVGLSPQSITDCLYIFTSGTGLPKARISHLKLQCGFYQ 360
DB 301 ISDLAEVSAEVDGVPVGLSPQSITDCLYIFTSGTGLPKARISHLKLQCGFYQ 360
QY 361 LCGVHOEDVLYALPLYNKSGSLGIVGCMGIGATVILKSFSAQFWEDECQHRVYFQ 420
DB 361 LCGVHOEDVLYALPLYNKSGSLGIVGCMGIGATVILKSFSAQFWEDECQHRVYFQ 420
QY 421 YIGELCRVYNOPPSKAERGHKVRILAVSGLRPDTWERFVRFGQLVLETYGTGNA 480
DB 421 YIGELCRVYNOPPSKAERGHKVRILAVSGLRPDTWERFVRFGQLVLETYGTGNA 480
QY 481 TINYTGORGAVRASMLYKHIPFSLIRYDVTGGPIRDPOGCMATSPGEGGLVAVS 540
DB 481 TINYTGORGAVRASMLYKHIPFSLIRYDVTGGPIRDPOGCMATSPGEGGLVAVS 540
QY 541 QOSPFLGAGGPELQGLKLDVFRPGVFFNTGSLVCDGQGFRRFDRGDTFRWGE 600
DB 541 QOSPFLGAGGPELQGLKLDVFRPGVFFNTGSLVCDGQGFRRFDRGDTFRWGE 600
QY 601 NVATTEVAEVEFALDFLOEVNYYGTVPGHEGRAGAAALVLRPFLALDMLQYTHVENL 660
DB 601 NVATTEVAEVEFALDFLOEVNYYGTVPGHEGRAGAAALVLRPFLALDMLQYTHVENL 660
QY 661 PYYARRPRLQESLATTITTFKQKVRANSGFDSITSDPLVYLDQVAGYLPITARY 720
DB 661 PYYARRPRLQESLATTITTFKQKVRANSGFDSITSDPLVYLDQVAGYLPITARY 720

QY 721 SALLAGNLRI 730
DB 721 SALLAGNLRI 730
RESULT 13
ABO19656
ID ABO19656 standard; protein; 730 AA.
XX
XX
AC ABO19656;
XX
DT 08-SEP-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO703.

XX Human, secreted and transmembrane protein; PRO, cell death; neuropathy;
XX peripheral neuropathy; diabetic peripheral neuropathy;
XX AIDS-associated neuropathy; Charcot-Marie-Tooth disease;
XX Refsum's disease; Abetalipoproteinemia; Tangier disease;
XX Krabbe's disease; Metachromatic leukodystrophy; Fabry's disease;
XX Dejerine-Sottas syndrome; chromosome mapping; gene mapping; gene therapy.
XX Homo sapiens.
XX
XX US2003050240-A1.
XX
PD 13-MAR-2003.
XX
PF 16-OCT-2001; 2001US-00978403.
XX
XX 17-OCT-1997; 97US-0062250P.
XX 03-NOV-1997; 97US-0064249P.
XX 13-NOV-1997; 97US-0065311P.
XX 21-NOV-1997; 97US-0066364P.
XX 10-MAR-1998; 98US-0077450P.
XX 11-MAR-1998; 98US-0077532P.
XX 11-MAR-1998; 98US-0077641P.
XX 11-MAR-1998; 98US-0077649P.
XX 12-MAR-1998; 98US-0077791P.
XX 13-MAR-1998; 98US-0078004P.
XX 20-MAR-1998; 98US-0078886P.
XX 20-MAR-1998; 98US-0078910P.
XX 20-MAR-1998; 98US-0078936P.
XX 20-MAR-1998; 98US-0078939P.
XX 26-MAR-1998; 98US-0079294P.
XX 26-MAR-1998; 98US-0079656P.
XX 27-MAR-1998; 98US-0079663P.
XX 27-MAR-1998; 98US-0079664P.
XX 27-MAR-1998; 98US-0079689P.
XX 27-MAR-1998; 98US-0079728P.
XX 27-MAR-1998; 98US-0079786P.
XX 30-MAR-1998; 98US-0079920P.
XX 30-MAR-1998; 98US-0079923P.
XX 31-MAR-1998; 98US-0080105P.
XX 31-MAR-1998; 98US-0080107P.
XX 31-MAR-1998; 98US-0080165P.
XX 31-MAR-1998; 98US-0080194P.
XX 01-APR-1998; 98US-0080327P.
XX 01-APR-1998; 98US-0080328P.
XX 01-APR-1998; 98US-0080333P.
XX 01-APR-1998; 98US-0080344P.
XX 08-APR-1998; 98US-0081049P.
XX 08-APR-1998; 98US-0081070P.
XX 08-APR-1998; 98US-0081071P.
XX 08-APR-1998; 98US-0081195P.
XX 09-APR-1998; 98US-0081203P.
XX 09-APR-1998; 98US-0081229P.
XX 15-APR-1998; 98US-0081817P.
XX 15-APR-1998; 98US-0081819P.
XX 15-APR-1998; 98US-0081838P.
XX 15-APR-1998; 98US-0081852P.
XX 15-APR-1998; 98US-0081955P.

```

PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082700P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 22-APR-1998; 98US-0082804P.
PR 23-APR-1998; 98US-0082796P.
PR 27-APR-1998; 98US-0083333P.
PR 28-APR-1998; 98US-0083332P.
PR 29-APR-1998; 98US-0083392P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083500P.
PR 29-APR-1998; 98US-0083545P.
PR 29-APR-1998; 98US-0083546P.
PR 29-APR-1998; 98US-0083558P.
PR 29-APR-1998; 98US-0083559P.
PR 30-APR-1998; 98US-0083742P.
PR 05-MAY-1998; 98US-0084366P.
PR 05-MAY-1998; 98US-0084414P.
PR 06-MAY-1998; 98US-0084441P.
PR 07-MAY-1998; 98US-0084598P.
PR 07-MAY-1998; 98US-0084600P.
PR 07-MAY-1998; 98US-0084627P.
PR 07-MAY-1998; 98US-0084637P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 13-MAY-1998; 98US-0085323P.
PR 13-MAY-1998; 98US-0085338P.
PR 15-MAY-1998; 98US-0085339P.
PR 15-MAY-1998; 98US-0085573P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085689P.
PR 15-MAY-1998; 98US-0085697P.
PR 15-MAY-1998; 98US-0085700P.
PR 15-MAY-1998; 98US-0085704P.
PR 18-MAY-1998; 98US-0086032P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086414P.
PR 22-MAY-1998; 98US-0086430P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087106P.
PR 28-MAY-1998; 98US-0087208P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 30-JUL-1998; 98US-0094651P.
PR 11-SEP-1998; 98US-0100038P.
PR 07-OCT-1998; 98US-0100211P.
PR 20-NOV-1998; 98US-0109304P.
PR 20-NOV-1998; 98US-0109304P.
PR 22-DEC-1998; 98US-0112266P.
PR 23-DEC-1998; 98US-0116212P.
PR 05-JAN-1999; 98US-0116212P.
PR 08-MAR-1999; 98US-0134287P.
PR 10-MAR-1999; 98US-0134287P.
PR 12-MAR-1999; 98US-0134287P.
PR 29-MAR-1999; 98US-0134287P.
PR 21-APR-1999; 98US-0134287P.
PR 26-APR-1999; 98US-0134452P.
PR 28-APR-1999; 98US-0134452P.
PR 14-MAY-1999; 98US-0134287P.
PR 02-JUN-1999; 98US-0134287P.
PR 16-JUN-1999; 98US-0134287P.
PR 23-JUN-1999; 98US-0134287P.
PR 07-JUL-1999; 98US-0142680P.
PR 26-JUL-1999; 98US-0145698P.

```

```

PR 28-JUL-1999; 98US-0146222P.
PR 28-OCT-1999; 98US-0162506P.
PR 30-NOV-1999; 98US-0502831P.
PR 02-DEC-1999; 98US-0502855P.
PR 02-DEC-1999; 98US-0502855P.
PR 16-DEC-1999; 98US-0503009P.
PR 30-DEC-1999; 98US-0503124P.
PR 05-JAN-1999; 98US-0503127P.
PR 05-JAN-2000; 98US-0500219P.
PR 06-JAN-2000; 98US-0500277P.
PR 06-JAN-2000; 98US-0500376P.
PR 11-FEB-2000; 98US-0500356P.
PR 18-FEB-2000; 98US-0500341P.
PR 24-FEB-2000; 98US-0500500P.
PR 02-MAR-2000; 98US-0500581P.
PR 10-MAR-2000; 98US-0500531P.
PR 21-MAR-2000; 98US-0500753P.
PR 30-MAR-2000; 98US-0500843P.
PR 17-MAY-2000; 98US-0501370P.
PR 22-MAY-2000; 98US-0501404P.
PR 30-MAY-2000; 98US-0501494P.
PR 02-JUN-2000; 98US-0501526P.
PR 28-JUL-2000; 98US-0502071P.
PR 28-AUG-2000; 98US-0502328P.
PR 01-DEC-2000; 98US-0503678P.
PR 20-DEC-2000; 98US-0503678P.
PR 28-FEB-2001; 98US-0506520P.
PR 22-MAR-2001; 98US-0509552P.
PR 25-MAY-2001; 98US-0501709P.
PR 01-JUN-2001; 98US-0501780P.
PR 20-JUN-2001; 98US-0501969P.
PR 29-JUN-2001; 98US-0502106P.
PR 09-JUL-2001; 98US-0502173P.
PR 30-JUL-2001; 98US-0502185P.
XX
XX
XX (GENTECH ) GENENTECH INC.
XX
XX Ashkenazi AJ, Baker KP, Botstein D, Deenoyers J, Eaton DL,
XX Ferrara N, Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME,
XX Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan MJ,
XX Kijavain IJ, Kuo SS, Napier MA, Pan J, Paoni NF, Roy MA, Shelton DL,
XX Stewart TA, Tumas D, Williams PM, Wood WL,
XX
XX WPI, 2003-503575/47.
XX N-PSDB; ACD29691.
XX
XX Novel secreted and transmembrane polypeptide for modulating biological
XX activity of cell expressing the polypeptide, identifying agonists or
XX antagonists of polypeptide, and as molecular weight markers.
XX
XX Claim 12; Fig 39; 459pp; English.
XX
XX
XX
XX The invention describes an isolated, secreted and transmembrane
XX polypeptide, termed PRO polypeptide (I). (I) is useful for detecting
XX PRO493, PRO337, PRO1559, PRO725, PRO700 or PRO739 polypeptide, and for
XX linking a bioactive molecule to a cell expressing the above polypeptide.
XX The bioactive molecule is a toxin, radiolabel or an antibody and causes
XX cell death. (I) is useful as therapeutic agent, in medical and industrial
XX applications e.g. for treating neuropathy, AIDS-associated neuropathy,
XX neuropathy, diabetic peripheral neuropathy, AIDS-associated neuropathy,
XX Charcot-Marie-Tooth disease, Refsum's disease, Abetalipoproteinemia,
XX Tangle disease, Krabbe's disease, Metachromatic leukodystrophy, Fabry's
XX

```

```

Query Match 100.0%; Score 3843; DB 6; Length 730;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MGVCORTPAFWKKSQLEBAALGFRKGSAGMFGNQTVPLEAGSMALLPILL 60
DB 1 MGVCORTPAFWKKSQLEBAALGFRKGSAGMFGNQTVPLEAGSMALLPILL 60
QY 61 PLLILKLIHMPQIRKMLPADIAFVRALCCGRALRAFAALAAAADPGEGGCSLMRLAE 120

```

```
Db 61 PLLIKLHIMPQLRLPADLAFVAVALCCKRLRLRALAAAAADPEEGCCSLAKRLAE 120
Qy 121 LAQRAAHTLLIHGSRFRFSYSAERESNRPAAPALPRLGMDGPDGDSGGSGAGEGERRA 180
Db 121 LAQRAAHTFLIHGSRFRFSYSAERESNRPAAPALPRLGMDGPDGDSGGSGAGEGERRA 180
Qy 181 ARGADDAAGSGAEFAGGAGGARGGAAPLSPGATVLLLLPAGEPEFLMWFLAKXGLR 240
Db 181 ARGADDAAGSGAEFAGGAGGARGGAAPLSPGATVLLLLPAGEPEFLMWFLAKXGLR 240
Qy 241 TAFVPTALRRGPLLHCLRSQCARALVLAPEFLSELPDLPALRAMGLHLMAGPQTIPAG 300
Db 241 TAFVPTALRRGPLLHCLRSQCARALVLAPEFLSELPDLPALRAMGLHLMAGPQTIPAG 300
Qy 301 ISDLIAEVAEVDGVPGLTSPGISTDTCTCYIFTSGTGPKARISHLTLKLOCGFYQ 360
Db 301 ISDLIAEVAEVDGVPGLTSPGISTDTCTCYIFTSGTGPKARISHLTLKLOCGFYQ 360
Qy 361 LCGVHOEDVLYLALPLYHMSGSLGI VGC MGIGATVVLKSKFSAGQFWECCQHRVTVFQ 420
Db 361 LCGVHOEDVLYLALPLYHMSGSLGI VGC MGIGATVVLKSKFSAGQFWECCQHRVTVFQ 420
Qy 421 YIGELCRYLVNPPSKAEKGRVLA VSGSLRPDTWEPVRRFGPLQVLTETYGLETEGVA 480
Db 421 YIGELCRYLVNPPSKAEKGRVLA VSGSLRPDTWEPVRRFGPLQVLTETYGLETEGVA 480
Qy 481 TINYGORGAVGRASMLYKHIFPFLIRYDTTGEPIHDPOGCMATSPGEPGLVA PVS 540
Db 481 TINYGORGAVGRASMLYKHIFPFLIRYDTTGEPIHDPOGCMATSPGEPGLVA PVS 540
Qy 541 QOSPFLIAGVGPPELLAQGKLKDVPRPGDVFNNTGDLVCDGGLRPFHDFRGTDFRWKGE 600
Db 541 QOSPFLIAGVGPPELLAQGKLKDVPRPGDVFNNTGDLVCDGGLRPFHDFRGTDFRWKGE 600
Qy 601 NVATTEVAEVEFALDFLOEVNYYGVTPGHEGRAGMALVLRPPALDLMQLYTHVSEN 660
Db 601 NVATTEVAEVEFALDFLOEVNYYGVTPGHEGRAGMALVLRPPALDLMQLYTHVSEN 660
Qy 661 PPYARPRFLRQESLATTETFKQOKVRMANEGFDDSTLSDPLYYLDQAVAYLPPLTARY 720
Db 661 PPYARPRFLRQESLATTETFKQOKVRMANEGFDDSTLSDPLYYLDQAVAYLPPLTARY 720
Qy 721 SALLAGNLRI 730
Db 721 SALLAGNLRI 730

RESULT 14
ADA12302
ID ADA12302 standard; protein; 730 AA.
XX
AC ADA12302;
XX
DT 06-NOV-2003 (first entry)
XX
DE Human secreted/transmembrane polypeptide PRO703.
XX
KW inflammatory disease; organ failure; atherosclerosis; cardiac injury;
KW infectility; birth defect; premature aging; AIDS; cancer;
KW diabetic complication; tissue typing; human.
XX
OS Homo sapiens.
XX
PN US2003055216-A1.
XX
PD 20-MAR-2003.
XX
PF 17-OCT-2001; 2001US-00978824.
XX
PR 21-MAY-1996; 96US-0018049P.
PR 17-OCT-1997; 97US-0062250P.
PR 03-NOV-1997; 97US-0064249P.
PR 13-NOV-1997; 97US-0065311P.
```

PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085689P.
PR 15-MAY-1998; 98US-0085697P.
PR 15-MAY-1998; 98US-0085700P.
PR 15-MAY-1998; 98US-0085704P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086414P.
PR 22-MAY-1998; 98US-0086430P.
PR 22-MAY-1998; 98US-0086488P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087106P.
PR 28-MAY-1998; 98US-0087208P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 30-JUL-1998; 98US-0094651P.
PR 11-SEP-1998; 98US-0100038P.
PR 07-OCT-1998; 98US-00168978.
PR 07-OCT-1998; 98US-00168978.
PR 02-NOV-1998; 98US-0021141.
PR 06-NOV-1998; 98US-00184216.
PR 06-NOV-1998; 98US-00187358.
PR 20-NOV-1998; 98US-0109304P.
PR 20-NOV-1998; 98US-0024855.
PR 07-DEC-1998; 98US-00202054.
PR 22-DEC-1998; 98US-00218517.
PR 23-DEC-1998; 98US-0113296P.
PR 05-JAN-1999; 98US-00200106.
PR 05-MAR-1999; 98US-00254455.
PR 08-MAR-1999; 98US-002505028.
PR 10-MAR-1999; 98US-00265686.
PR 10-MAR-1999; 98US-00265686.
PR 12-MAR-1999; 98US-00267213.
PR 12-MAR-1999; 98US-0123957P.
PR 29-MAR-1999; 98US-01267737P.
PR 12-APR-1999; 98US-00284281.
PR 21-APR-1999; 98US-0130232P.
PR 26-APR-1999; 98US-0131022P.
PR 28-APR-1999; 98US-00314452.
PR 14-MAY-1999; 98US-0134287P.
PR 14-MAY-1999; 98US-0134287P.
PR 02-JUN-1999; 98US-00310733.
PR 16-JUN-1999; 98US-00310733.
PR 23-JUN-1999; 98US-0139557P.
PR 07-JUL-1999; 98US-0141037P.
PR 26-JUL-1999; 98US-014680P.
PR 28-JUL-1999; 98US-0146223P.
PR 25-AUG-1999; 98US-00380137.
PR 25-AUG-1999; 98US-00380138.
PR 25-AUG-1999; 98US-00380142.
PR 29-OCT-1999; 98US-0162506P.
PR 30-NOV-1999; 98US-0028313.
PR 02-DEC-1999; 98US-0028551.
PR 02-DEC-1999; 98US-0028551.
PR 16-DEC-1999; 98US-0030955.
PR 30-DEC-1999; 98US-0031243.
PR 30-DEC-1999; 98US-0031274.
PR 05-JAN-2000; 2000US-0000217.
PR 06-JAN-2000; 2000US-0000277.
PR 11-FEB-2000; 2000US-0003565.
PR 18-FEB-2000; 2000US-0003431.
PR 24-FEB-2000; 2000US-0005004.
PR 02-MAR-2000; 2000US-0005841.
PR 10-MAR-2000; 2000US-0006319.
PR 21-MAR-2000; 2000US-0007532.
PR 30-MAR-2000; 2000US-0008439.
PR 17-MAY-2000; 2000US-0013705.
PR 22-MAY-2000; 2000US-0014042.

PR 30-MAY-2000; 2000US-0014941.
PR 02-JUN-2000; 2000US-0015264.
PR 28-JUL-2000; 2000US-0020710.
PR 24-AUG-2000; 2000US-0023328.
PR 08-NOV-2000; 2000US-00709238.
PR 27-NOV-2000; 2000US-00723749.
PR 01-DEC-2000; 2000US-00742678.
PR 20-DEC-2000; 2000US-0074259.
PR 20-DEC-2000; 2000US-0034956.
PR 28-FEB-2001; 2001US-0086520.
PR 22-MAR-2001; 2001US-00815744.
PR 22-MAR-2001; 2001US-00816920.
PR 22-MAR-2001; 2001US-00816952.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854208.
PR 21-MAY-2001; 2001US-0017092.
PR 01-JUN-2001; 2001US-00872035.
PR 01-JUN-2001; 2001US-0017800.
PR 05-JUN-2001; 2001US-00874503.
PR 14-JUN-2001; 2001US-00882636.
PR 19-JUN-2001; 2001US-00886342.
PR 20-JUN-2001; 2001US-0019692.
PR 29-JUL-2001; 2001US-0021735.
PR 09-JUL-2001; 2001US-0021735.
PR 30-JUL-2001; 2001US-00918585.

XX (GETH) GENENTECH INC.
PA Ashkenazi AJ, Baker KP, Botstein D, Deeneyers L, Eaton DJ,
XX Pi Ferrara N, Filvarcoff E, Fong S, Gao W, Gerber H, Gertlisen ME;
PI

Query Match 100.0%; Score 3843; DB 6; Length 730;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVCORTRAPMKKESQLEBAALGFRKSGSGMFASGNNQVPIEAGSMALLPLLL 60
DB 1 MGVCORTRAPMKKESQLEBAALGFRKSGSGMFASGNNQVPIEAGSMALLPLLL 60

QY PLLLLKLHMPQRLMPLADLAFVRLCKRLRARALAAAADPBGCGCSLARLAE 120
DB PLLLLKLHMPQRLMPLADLAFVRLCKRLRARALAAAADPBGCGCSLARLAE 120

QY 61 PLLLLKLHMPQRLMPLADLAFVRLCKRLRARALAAAADPBGCGCSLARLAE 120
DB 61 PLLLLKLHMPQRLMPLADLAFVRLCKRLRARALAAAADPBGCGCSLARLAE 120

QY 121 LAQORAAHFTLHSGRFSYSEABERESNRAPALFALMDGPDGDSGSGAGCERA 180
DB 121 LAQORAAHFTLHSGRFSYSEABERESNRAPALFALMDGPDGDSGSGAGCERA 180

QY 181 APGAGDAAAAGSAGFAGDGAARGGGAAALPSGATVALLPAPGFILWFGLAAGLR 240
DB 181 APGAGDAAAAGSAGFAGDGAARGGGAAALPSGATVALLPAPGFILWFGLAAGLR 240

QY 241 TAFVPTALRRGPLHCLRSOGARALVLAPEFLESLEPDI PALRAWGLHMAAGPETHAG 300
DB 241 TAFVPTALRRGPLHCLRSOGARALVLAPEFLESLEPDI PALRAWGLHMAAGPETHAG 300

QY 301 ISDLAEVSAEVDGPPVGYLSPQSTITDTCIVYFTSGTGLPXAARISHKLILQCCGFYQ 360
DB 301 ISDLAEVSAEVDGPPVGYLSPQSTITDTCIVYFTSGTGLPXAARISHKLILQCCGFYQ 360

QY 361 LCGVHQBVDYIYALPLYHMSGSLIGVCGMGATVYLSKKSAGFMEDCQHRVTYQ 420
DB 361 LCGVHQBVDYIYALPLYHMSGSLIGVCGMGATVYLSKKSAGFMEDCQHRVTYQ 420

QY 421 YIGELCRYVNOPPSAERGHKRVLAWSGLRPDTERFVRRGPQVLETYGLTEGNYA 480
DB 421 YIGELCRYVNOPPSAERGHKRVLAWSGLRPDTERFVRRGPQVLETYGLTEGNYA 480

QY 481 TINYTGQRAVGSASWLYKHIPPFLIRYDVTTGEPDIRPDQHCVAATSGEPGLVAPVS 540
DB 481 TINYTGQRAVGSASWLYKHIPPFLIRYDVTTGEPDIRPDQHCVAATSGEPGLVAPVS 540

QY 541 QOSPFLGYAGGPELAAQKLLKQVFRPGDVPFNTGDLVCCDDGFLRFHRTGDPFKWGE 600
DB 541 QOSPFLGYAGGPELAAQKLLKQVFRPGDVPFNTGDLVCCDDGFLRFHRTGDPFKWGE 600

Db 541 QOSPFLGVAGPELDAQKLIKDVFRPGDVFFNTGDLIVCDQGFRLFRHRTGDTFRWKGE 600
Qy NATTEVEVEFEFALDFOEYVNVYGVTPGHEGRAMALVVRPPALDLMOLYHVSBNL 660
Db 601 NATTEVEVEFEFALDFOEYVNVYGVTPGHEGRAMALVVRPPALDLMOLYHVSBNL 660
Qy 661 PVPARPRFLRLOESLATTETFPKQKVVRMANEGFDPSTLSDPLVLDQAVGAYLPLTARY 720
Db 661 PVPARPRFLRLOESLATTETFPKQKVVRMANEGFDPSTLSDPLVLDQAVGAYLPLTARY 720
Qy 721 SALLAGNLR 730
Db 721 SALLAGNLR 730
RESULT 15
ABO19547
ID ABO19547 standard; protein; 730 AA.
AC ABO19547;
XX
XX 27-AUG-2003 (first entry)
DE Novel human secreted and transmembrane polypeptide #15.
XX
XX Human; secreted and transmembrane protein; PRO; viral infection;
KW tumour growth; retinal disorder; injury; sight loss;
KW retinitis pigmentosa; age-related macular degeneration;
KW sport-related joint problem; articular cartilage defect; osteoarthritis;
KW rheumatoid arthritis; wound healing; obesity; diabetes; insulinemia;
KW kidney disorder; mesangial cell function; Berger disease; nephropathy;
KW cardiac disease; dermatitis; Crohn disease; neuropathy;
KW diabetic peripheral neuropathy; autonomic neuropathy;
KW reduced motility of the gastrointestinal tract;
KW atony of the urinary bladder; post polio syndrome; Krabbe's disease;
KW Charcot-Marie-Tooth disease; Fabry's disease; Tanager disease;
KW Refsum's disease.
XX
XX Homo sapiens.
OS
XX
XX US2003049633-A1.
PD 13-MAR-2003.
XX
PF 16-OCT-2001; 2001US-00978585.
XX
XX 17-OCT-1997; 97US-0062250P.
PR 03-NOV-1997; 97US-0064249P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066364P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077641P.
PR 11-MAR-1998; 98US-0077649P.
PR 12-MAR-1998; 98US-0077791P.
PR 13-MAR-1998; 98US-0078004P.
PR 17-MAR-1998; 98US-0080422P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078910P.
PR 20-MAR-1998; 98US-0078936P.
PR 20-MAR-1998; 98US-0078939P.
PR 25-MAR-1998; 98US-0079294P.
PR 26-MAR-1998; 98US-0079656P.
PR 27-MAR-1998; 98US-0079663P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079689P.
PR 27-MAR-1998; 98US-0079728P.
PR 27-MAR-1998; 98US-0079786P.
PR 30-MAR-1998; 98US-0079920P.
PR 30-MAR-1998; 98US-0079923P.
PR 31-MAR-1998; 98US-0080105P.
PR 31-MAR-1998; 98US-0080107P.

PR 31-MAR-1998; 98US-0080165P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080328P.
PR 01-APR-1998; 98US-0080332P.
PR 01-APR-1998; 98US-0080334P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 08-APR-1998; 98US-0081071P.
PR 09-APR-1998; 98US-0081195P.
PR 09-APR-1998; 98US-0081203P.
PR 09-APR-1998; 98US-0081229P.
PR 15-APR-1998; 98US-0081817P.
PR 15-APR-1998; 98US-0081819P.
PR 15-APR-1998; 98US-0081838P.
PR 15-APR-1998; 98US-0081952P.
PR 15-APR-1998; 98US-0081955P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082700P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 22-APR-1998; 98US-0082804P.
PR 23-APR-1998; 98US-0082796P.
PR 27-APR-1998; 98US-0083336P.
PR 28-APR-1998; 98US-0083322P.
PR 28-APR-1998; 98US-0083392P.
PR 28-APR-1998; 98US-0083495P.
PR 28-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083500P.
PR 29-APR-1998; 98US-0083545P.
PR 29-APR-1998; 98US-0083554P.
PR 29-APR-1998; 98US-0083558P.
PR 29-APR-1998; 98US-0083559P.
PR 30-APR-1998; 98US-0083742P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 06-MAY-1998; 98US-0084441P.
PR 07-MAY-1998; 98US-0084598P.
PR 07-MAY-1998; 98US-0084600P.
PR 07-MAY-1998; 98US-0084627P.
PR 07-MAY-1998; 98US-0084637P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 13-MAY-1998; 98US-0085323P.
PR 13-MAY-1998; 98US-0085338P.
PR 13-MAY-1998; 98US-0085339P.
PR 15-MAY-1998; 98US-0085573P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085800P.
PR 15-MAY-1998; 98US-0085822P.
PR 15-MAY-1998; 98US-0085892P.
PR 15-MAY-1998; 98US-0085977P.
PR 15-MAY-1998; 98US-0085700P.
PR 15-MAY-1998; 98US-0085704P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086414P.
PR 22-MAY-1998; 98US-0086430P.
PR 22-MAY-1998; 98US-0086486P.
PR 22-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087099P.
PR 28-MAY-1998; 98US-0087106P.
PR 28-MAY-1998; 98US-0087208P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 30-JUL-1998; 98US-0094651P.
PR 11-SEP-1998; 98US-0100038P.
PR 07-OCT-1998; 98US-00168978.
PR 07-OCT-1998; 98WO-US021141.

PR 02-NOV-1998; 98US-00184216.
PR 06-NOV-1998; 98US-00187368.
PR 20-NOV-1998; 98US-0109304P.
PR 20-NOV-1998; 98MO-US024855.
PR 07-DEC-1998; 98US-00202054.
PR 22-DEC-1998; 98US-00218517.
PR 23-DEC-1998; 98US-0113296P.
PR 22-DEC-1998; 98US-0113621P.
PR 05-JAN-1999; 99MO-US000106.
PR 05-MAR-1999; 99US-00254465.
PR 08-MAR-1999; 99MO-US005028.
PR 10-MAR-1999; 99US-00255686.
PR 12-MAR-1999; 99MO-US005190.
PR 12-MAR-1999; 99US-00267213.
PR 29-MAR-1999; 99US-0126773P.
PR 12-APR-1999; 99US-00284291.
PR 21-APR-1999; 99US-0130232P.
PR 26-APR-1999; 99US-0131022P.
PR 28-APR-1999; 99US-0131445P.
PR 14-MAY-1999; 99US-00311832.
PR 14-MAY-1999; 99US-0134287P.
PR 02-JUN-1999; 99MO-US010733.
PR 16-JUN-1999; 99US-0139557P.
PR 23-JUN-1999; 99US-0141037P.
PR 07-JUL-1999; 99US-0142680P.
PR 26-JUL-1999; 99US-0145698P.
PR 28-JUL-1999; 99US-0146222P.
PR 25-AUG-1999; 99US-00380137.
PR 25-AUG-1999; 99US-00380138.
PR 29-OCT-1999; 99US-0162506P.
PR 30-NOV-1999; 99MO-US028313.
PR 02-DEC-1999; 99MO-US028551.
PR 16-DEC-1999; 99MO-US030095.
PR 30-DEC-1999; 99MO-US031243.
PR 30-DEC-1999; 99MO-US031274.
PR 05-JAN-2000; 2000MO-US000219.
PR 06-JAN-2000; 2000MO-US000277.
PR 11-FEB-2000; 2000MO-US003565.
PR 18-FEB-2000; 2000MO-US004341.
PR 24-FEB-2000; 2000MO-US005004.
PR 02-MAR-2000; 2000MO-US005841.
PR 10-MAR-2000; 2000MO-US006319.
PR 21-MAR-2000; 2000MO-US007532.
PR 30-MAR-2000; 2000MO-US008439.
PR 17-MAY-2000; 2000MO-US013705.
PR 22-MAY-2000; 2000MO-US014042.
PR 30-MAY-2000; 2000MO-US014941.
PR 02-JUN-2000; 2000MO-US015254.
PR 28-JUL-2000; 2000MO-US020710.
PR 24-AUG-2000; 2000MO-US023328.
PR 08-NOV-2000; 2000US-00709238.
PR 27-NOV-2000; 2000US-00723749.
PR 01-DEC-2000; 2000MO-US032678.
PR 20-DEC-2000; 2000US-00747259.
PR 20-DEC-2000; 2000MO-US034956.
PR 28-FEB-2001; 2001MO-US006520.
PR 22-MAR-2001; 2001US-00816744.
PR 22-MAR-2001; 2001US-00816920.
PR 22-MAR-2001; 2001MO-US009552.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 25-MAY-2001; 2001MO-US017092.
PR 01-JUN-2001; 2001US-00872035.
PR 01-JUN-2001; 2001MO-US017800.
PR 05-JUN-2001; 2001US-00874503.
PR 14-JUN-2001; 2001US-00882636.

Query Match 100.0%; Score 3843; DB 6; Length 730;

	Best Local Similarity	100.0%;	Pred. No. 0;	Mismatches	730;	Conservative	0;	Indels	0;	Gaps	0;
QY	1	MGVCORTRAPWKEKSLQJERAAALGFRKSGSFGWNOVPIEAGSMALLPLILL	60								
Db	1	MGVCORTRAPWKEKSLQJERAAALGFRKSGSFGWNOVPIEAGSMALLPLILL	60								
QY	61	PLILLKLHLPQJRMPLADLAFVRALCKKRALRABALAAAADPEPGGCSLANRLAE	120								
Db	61	PLILLKLHLPQJRMPLADLAFVRALCKKRALRABALAAAADPEPGGCSLANRLAE	120								
QY	121	LAQQRRAHTFLIHGSRFSYSSEAEBSNRARAFRALGMDGPDGDSGSGSAGEGERA	180								
Db	121	LAQQRRAHTFLIHGSRFSYSSEAEBSNRARAFRALGMDGPDGDSGSGSAGEGERA	180								
QY	181	APGAGDAAAGSAGFAFGDGAARCGGAAAPLSPGATVALLPAGPEFLMFGLAQAGUR	240								
Db	181	APGAGDAAAGSAGFAFGDGAARCGGAAAPLSPGATVALLPAGPEFLMFGLAQAGUR	240								
QY	241	TAFVPTALRRGPLHLCRSCGARALVLAPEFLESLEPDIPLARMGLHMAAGPETHPAG	300								
Db	241	TAFVPTALRRGPLHLCRSCGARALVLAPEFLESLEPDIPLARMGLHMAAGPETHPAG	300								
QY	301	ISDLAEVSAEVDGVPVGYLSSPOSITDTCLYIFTSGTGLPKAARISHKLIOCGFYQ	360								
Db	301	ISDLAEVSAEVDGVPVGYLSSPOSITDTCLYIFTSGTGLPKAARISHKLIOCGFYQ	360								
QY	361	LCGVHGEDVIYIALPLYHNGSILGIVGCMGIGATVVLKSKTSAGOFMEDCOOHRVTYQ	420								
Db	361	LCGVHGEDVIYIALPLYHNGSILGIVGCMGIGATVVLKSKTSAGOFMEDCOOHRVTYQ	420								
QY	421	YIGELCRYLVNOPSKAERGHKRYRLAVSGSLRPTDERFVRFPGLQVETYGLTGEGNVA	480								
Db	421	YIGELCRYLVNOPSKAERGHKRYRLAVSGSLRPTDERFVRFPGLQVETYGLTGEGNVA	480								
QY	481	TINYTGORGAVRASWLYXHIFFPSLIRYDVTGEBIRDPQHCMAISGEPGLVAPVS	540								
Db	481	TINYTGORGAVRASWLYXHIFFPSLIRYDVTGEBIRDPQHCMAISGEPGLVAPVS	540								
QY	541	QOSPFLGYAGPPELLAGKULXDFRPGDVFPMTGDLVCDGDFLRFHRTGDTFFPKGE	600								
Db	541	QOSPFLGYAGPPELLAGKULXDFRPGDVFPMTGDLVCDGDFLRFHRTGDTFFPKGE	600								
QY	601	NVATTEVAEVEFALDLOEVNNGVTVPGHEGRAGYALVLRPHALDLMQYTHVSENL	660								
Db	601	NVATTEVAEVEFALDLOEVNNGVTVPGHEGRAGYALVLRPHALDLMQYTHVSENL	660								
QY	661	PPYARPPFLRLQSLATTETPFKQKRYRANEGFDBTSLDPLVLDQAVGAYLPLTTARY	720								
Db	661	PPYARPPFLRLQSLATTETPFKQKRYRANEGFDBTSLDPLVLDQAVGAYLPLTTARY	720								
QY	721	SALLAGNLR 730									
Db	721	SALLAGNLR 730									

Search completed: June 30, 2004, 17:30:37
Job time : 65 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2004, 17:30:45 ; Search time 52 Seconds

(without alignments)
4429.390 Million cell updates/sec

Title: US-10-030-226-2

Sequence: 1 MGVCGRTRAPWKEKSQLERA.....AYLPITRYSALLAGLRLI 730

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database: SPTREMBL_25:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_rhnc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_rudent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_virus:*

16: sp_bacteriaph:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3843	100.0	730	4	Q9BTY5
2	3683	95.8	700	4	Q9BTY5 homo sapien
3	2763	71.9	700	4	Q8N2X7
4	2103	54.7	446	11	Q88561
5	2083.5	54.2	402	4	Q8BTU5
6	1407	36.6	288	4	Q8TEU0
7	1353.5	35.2	619	4	Q7Z6E6
8	1350.5	35.1	619	4	Q86YF6
9	1346	35.0	620	11	Q9Y2P4
10	1346	34.2	625	13	Q8AVC5
11	1312.5	33.9	690	4	Q8Y2P5
12	1303.5	32.6	662	11	Q88694
13	1251	32.6	689	11	Q91VD5
14	1231	32.0	690	11	Q9E838
15	1193.5	31.1	623	11	Q88560

17	1076	28.0	638	16	Q89GR0	Q89GR0 bradyrhizob
18	1048.5	27.3	643	11	Q91VE0	Q91VE0 mus musculu
19	1043	27.1	604	16	Q89CL2	Q89CL2 bradyrhizob
20	1028	26.7	635	16	Q9A5Z8	Q9A5Z8 caulobacter
21	995	25.9	687	5	Q8SKR7	Q8SKR7 drosophila
22	992.5	25.8	661	5	Q9W185	Q9W185 drosophila
23	985.5	25.6	506	11	Q88562	Q88562 mus musculu
24	952	24.8	626	5	Q9YKUI	Q9YKUI drosophila
25	942.5	24.5	641	4	Q95186	Q95186 homo sapien
26	930.5	24.2	671	5	Q9W1W0	Q9W1W0 drosophila
27	924.5	24.1	679	5	Q811F9	Q811F9 drosophila
28	903	23.5	608	16	Q8HZV4	Q8HZV4 pseudomonas
29	859.5	22.4	464	5	Q8SKT7	Q8SKT7 drosophila
30	858	22.3	655	5	Q18916	Q18916 caenorhabdi
31	841.5	21.9	597	16	Q05307	Q05307 mycobacteri
32	841.5	21.9	597	16	Q7U0E2	Q7U0E2 mycobacteri
33	839.5	21.8	650	5	Q18878	Q18878 caenorhabdi
34	763.5	19.9	334	11	Q8CV67	Q8CV67 mus musculu
35	725.5	18.9	643	3	Q42633	Q42633 cochlilobu
36	711	18.5	609	3	Q8J0E9	Q8J0E9 cephalospor
37	669	17.4	669	3	Q6J021	Q6J021 saccharomyc
38	540.5	14.1	258	4	Q9BTX1	Q9BTX1 homo sapien
39	469.5	12.2	561	16	Q7W6V1	Q7W6V1 bordecella
40	466.5	12.1	561	16	Q7W6V5	Q7W6V5 bordecella
41	453	11.8	554	16	Q89PP7	Q89PP7 bradyrhizob
42	411	10.7	559	16	Q7W037	Q7W037 bordecella
43	406.5	10.6	522	16	Q83MG9	Q83MG9 shigella fl
44	403	10.5	502	16	Q7W6M7	Q7W6M7 bordecella
45	400	10.4	502	16	Q7W6N8	Q7W6N8 bordecella

ALIGNMENTS

RESULT 1

AC Q9BTY5; Q968W5; PRELIMINARY; PRT; 730 AA.

DT 01-JUN-2001 (TEMBLrel. 17, Created)

DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)

DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)

DE Hypothetical protein FLJ14593 (Hypothetical protein N72RP200142).

DE Homo sapiens (Human).

OS Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi; OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxId=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Lung;

RA Strausberg R.;

RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE OF 31-730 FROM N.A.

RA Isogai T., Ota T., Hayaishi K., Sugiyama T., Otsuki T., Suzuki Y., Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H., Nagatsuna M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M., Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S., Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y., Ninomiya K., Iwanaghi T.;

RT "NED human cDNA sequencing project";

RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RA Ota T., Nishikawa T., Suzuki Y., Kawai-Hio Y., Hayaishi K., Ishii S., Saito K., Yamamoto J., Wakamatsu A., Nagai T., Nakamura Y., Nagahara K., Sugano S., Isogai T.;

RT "HRI human cDNA sequencing project";

RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC003041; AAH03041.1; -

DR EMBL; BC009916; AAH09916.1; -

DR EMBL; AK027499; BAB55156.1; -

DR EMBL; AK075377; BAC11578.1; --
 DR Genew; HGNC:10997; SLC27A3.
 DR GO; GO:0003824; F:catalytic activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR000873; AMP-bind.
 DR Pfam; PF00501; AMP-binding; 2.
 DR PROSITE; PS00455; AMP-BINDING; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 730 AA; 78644 MW; E8CC4E9463A534A0 CRC64;

Query Match 100.0%; Score 3843; DB 4; Length 730;
 Best Local Similarity 100.0%; Pred. No. 1, 1e-255;
 Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NGVCORTRAPEKESQLEBALGFRKSGSWMFASGNTVTVEEGSMAALLPLLL 60
 DB 1 NGVCORTRAPEKESQLEBALGFRKSGSWMFASGNTVTVEEGSMAALLPLLL 60
 QY 61 PLLIKLHMPOLRLPADLFAVRLCCRALRARAALAAADPEGEGCSLAMP 120
 DB 61 PLLIKLHMPOLRLPADLFAVRLCCRALRARAALAAADPEGEGCSLAMP 120
 QY 121 LAQQAATFLIHGSRRESYSSEARESNRAAFALGMDWGPDGSGSAGEG 180
 DB 121 LAQQAATFLIHGSRRESYSSEARESNRAAFALGMDWGPDGSGSAGEG 180
 QY 181 AAGADDAAGSAGEAGDGAAGGGAAPISPGATVALLPAGPEFLMFGAKGLR 240
 DB 181 AAGADDAAGSAGEAGDGAAGGGAAPISPGATVALLPAGPEFLMFGAKGLR 240
 QY 241 TAFVPTALRGRPLHCLSCGALVLAPEFLESLEPDLPALRAMGLTMAAGGT 300
 DB 241 TAFVPTALRGRPLHCLSCGALVLAPEFLESLEPDLPALRAMGLTMAAGGT 300
 QY 301 ISDLAEVSAEVDGVPGLSSPGITDTCLYTFSSGTTGLPKARISHLILCO 360
 DB 301 ISDLAEVSAEVDGVPGLSSPGITDTCLYTFSSGTTGLPKARISHLILCO 360
 QY 361 LCGVHOEDVLYALPLHMSGLIGIVGCMGIGATVILKSKFSAGQFWDQ 420
 DB 361 LCGVHOEDVLYALPLHMSGLIGIVGCMGIGATVILKSKFSAGQFWDQ 420
 QY 421 YIGELCRYLVNPPSKAERGHKRLAVSGSLPDTWEPVRRFGLQVLE 480
 DB 421 YIGELCRYLVNPPSKAERGHKRLAVSGSLPDTWEPVRRFGLQVLE 480
 QY 481 TINYTGQGAAGVRAKMLKHLFPFSLIRYDTTGERPIIDPOGHCA 540
 DB 481 TINYTGQGAAGVRAKMLKHLFPFSLIRYDTTGERPIIDPOGHCA 540
 QY 541 QGSPFLGAGBELAQGLKLDVFRPQDVFNVTGDLVCDQGLRFH 600
 DB 541 QGSPFLGAGBELAQGLKLDVFRPQDVFNVTGDLVCDQGLRFH 600
 QY 601 NATTEVAVFEPALDFLOEVNNGVTVPGHSGRAMALVLRPPALD 660
 DB 601 NATTEVAVFEPALDFLOEVNNGVTVPGHSGRAMALVLRPPALD 660
 QY 661 PYYARPRFLRLOESLATTETFKQKVRMANEGFDSSTLDPYLDQAV 720
 DB 661 PYYARPRFLRLOESLATTETFKQKVRMANEGFDSSTLDPYLDQAV 720
 QY 721 SALLAGNLRI 730
 DB 721 SALLAGNLRI 730

RESULT 2

ID Q8N2X7 PRELIMINARY; PRT: 700 AA.
 AC Q8N2X7, 01-OCT-2002 (Tremblrel. 22, Created)
 DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)

DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 NX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (May-2002) to the EMBL/Genbank/DBJ databases.
 DR EMBL; BC029792; AAH29792.1; --
 DR GO; GO:0003824; F:catalytic activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR000873; AMP-bind.
 DR Pfam; PF00501; AMP-binding; 2.
 DR PROSITE; PS00455; AMP-BINDING; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 700 AA; 75356 MW; 4BE80A1D872A51A CRC64;

Query Match 95.8%; Score 3683; DB 4; Length 700;
 Best Local Similarity 100.0%; Pred. No. 1e-244;
 Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 MEASGNTQVPIEAGSMAALLPLLLPLLLKHLMPOLRLPADLFAVRLCC 90
 DB 1 MEASGNTQVPIEAGSMAALLPLLLPLLLKHLMPOLRLPADLFAVRLCC 90
 QY 91 PALRALAAAADPEGEGCSLAMPRLAEALQQAATFLIHGSRRESYS 150
 DB 91 PALRALAAAADPEGEGCSLAMPRLAEALQQAATFLIHGSRRESYS 150
 QY 151 ARAFLALGMDWGPDDGDSGEGSAGEGERAAGDAAAGSAGEAGDGA 210
 DB 151 ARAFLALGMDWGPDDGDSGEGSAGEGERAAGDAAAGSAGEAGDGA 210
 QY 211 LSPGATVALLPAGPEFLMFGAKGLRTAFVPTALRGRPLHCLSCG 270
 DB 211 LSPGATVALLPAGPEFLMFGAKGLRTAFVPTALRGRPLHCLSCG 270
 QY 271 FLESLEPDLPALRAMGLTMAAGGTTHPAGISDLAEVSAEVDGVP 330
 DB 271 FLESLEPDLPALRAMGLTMAAGGTTHPAGISDLAEVSAEVDGVP 330
 QY 331 LYFTSGTGLPKARISHLILCOGFFQLCGVHOEDVLYALPLHMSGL 390
 DB 331 LYFTSGTGLPKARISHLILCOGFFQLCGVHOEDVLYALPLHMSGL 390
 QY 391 GIGATVILKSKFSAGQFWDQCOHRTVFOYIGELCRYLVNPPSKA 450
 DB 391 GIGATVILKSKFSAGQFWDQCOHRTVFOYIGELCRYLVNPPSKA 450
 QY 451 LREDTWERFVRAPGLQVLETVGLTEGNVATINYTGQGAAGVRA 510
 DB 451 LREDTWERFVRAPGLQVLETVGLTEGNVATINYTGQGAAGVRA 510
 QY 511 VTTGERIPDPOGHCA 570
 DB 511 VTTGERIPDPOGHCA 570
 QY 571 FNTGDLVCDQGLRFH 630
 DB 571 FNTGDLVCDQGLRFH 630
 QY 631 EGRAGMAALVLRPPALDMLQVLYTHSENLPYARPRFLRLOESL 690
 DB 631 EGRAGMAALVLRPPALDMLQVLYTHSENLPYARPRFLRLOESL 690
 QY 691 EGFDPSLSDPLVLDQAVGAYLPTTARYSALLAGNLRI 730
 DB 691 EGFDPSLSDPLVLDQAVGAYLPTTARYSALLAGNLRI 730

RESULT 3

088561 PRELIMINARY; PRT; 614 AA.

AC 088561; (TREMblrel. 08, Created)

DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)

DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)

DE Fatty acid transport protein 3 (FATP3) (long-chain fatty acid transport protein 3) (Fragment).

GN SLC27A3.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_TaxId=10090;

RN (1)

RP SEQUENCE FROM N.A.

RX MEDLINE=98337965; PubMed=9671728;

RA Hirsch D., Stahl A., Lodish H.F.;

RT "A family of fatty acid transporters conserved from mycobacterium to man."

RL Proc. Natl. Acad. Sci. U.S.A. 95:8625-8629(1998).

CC -1- FUNCTION: INVOLVED IN TRANSLLOCATION OF LONG-CHAIN FATTY ACIDS ACROSS THE PLASMA MEMBRANE. MAY PLAY A PIVOTAL ROLE IN REGULATING AVAILABLE LONG-CHAIN FATTY ACID SUBSTRATES FROM EXOGENOUS SOURCES IN TISSUES UNDERGOING HIGH LEVELS OF BETA-OXIDATION OR TRIGLYCERIDE SYNTHESIS.

CC -1- TISSUE SPECIFICITY: LUNG, LIVER, AND TESTIS.

CC -1- SIMILARITY: TO OTHER ENZYMES WHICH ACT VIA AN ATP-DEPENDENT COVALENT BINDING OF AMP TO THEIR SUBSTRATE.

DR EMBL, AF072758; AAC40187.1; -.

DR MGD; MGI:1347358; SLC27A3.

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0003824; F:catalytic activity; IEA.

DR GO; GO:0008689; F:lipid transport; IEA.

DR GO; GO:0008152; P:metabolism; IEA.

DR InterPro; IPR000873; AMP-bind.

DR Pfam; PF00501; AMP-binding; 2.

DR PROSITE; PS00455; AMP BINDING; 1.

DR GlycoProtein; Lipid transport; Transmembrane; Transport.

FT NON TER 1 1

FT TRANSMEM 99 119 POTENTIAL.

FT TRANSMEM 262 282 POTENTIAL.

FT CARBOHYD 367 367 N-LINKED (GLCNAC. . .) (POTENTIAL).

SO SEQUENCE 614 AA; 67041 MW; 3302A558CDD9D989 CRC64;

Query Match

Best Local Similarity 71.9%; Score 2763; DB 11; Length 614;

Matches 526; Conservative 35; Mismatches 53; Indels 16; Gaps 2;

QY 101 AAADPEGEGGCSLMLRLAELAQCPAAFTFLIHGSRFRSSEAEERNSRRARAFALRGW 160

DB 1 AAADPESESGCSLMLRLAELAREQPTHFLIHGQRFSVAEAESESRRIRARALRGW 60

QY 161 DMGPFGSGSGSBAEGERAAPAGDAAAGSAEPAGDGAARGGGAAPISPGATVAL 220

DB 61 TGGRRG--SGRSTEGARVAPPADAAA-----RRTTAPPLAPGAYVAL 104

QY 221 LPAGEFTLMLWGKAGKRTAFVPTALRPGFLHCLRS CGARLVLAPEFLESLEP 280

DB 105 LPAGDFLIMWGLAKAGRTAFVPTALRPGFLHCLRS CGASALVLAPEFLESLEP 164

QY 281 ALRANGHLMAAGPETHPAGISDLIAEVSABVDGVPVGYLSSPGISDTCTIYFTSGTTG 340

DB 165 ALRANGHLMAAGPETHPAGISDLIAEVSABVDGVPVGYLSSPGISDTCTIYFTSGTTG 224

QY 341 LFKARISHLKLLOCGFYOLCGVHOEDVYIALPLVHMSGSLGIYVCMGIGATVVLKS 400

DB 225 LFKARISHLKLLOCGFYOLCGVHOEDVYIALPLVHMSGSLGIYVCMGIGATVVLKP 284

QY 401 KFSAGQFEMDCQGHVTVFOYIGELCRYLVNPPSKAERGHKVRILAVSGLRPPTWERFV 460

DB 285 KFSAGQFEMDCQGHVTVFOYIGELCRYLVNPPSKAERGHKVRILAVSGLRPPTWERFL 344

RESULT 4

088K70 PRELIMINARY; PRT; 446 AA.

AC 088K70;

DT 01-MAR-2003 (TREMblrel. 23, Created)

DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)

DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)

DE Solute carrier family 27.

GN SLC27A3.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_TaxId=10090;

RN (1)

RP SEQUENCE FROM N.A.

RX STRAIN=C57BL/6J; TISSUE=Body;

RX MEDLINE=42354683; PubMed=12466851;

RA The RIKEN Genome Consortium

RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs."

RL Nature 420:563-573(2002).

DR EMBL; AK076014; BAC36120.1; -.

DR MGD; MGI:1347358; SLC27A3.

DR GO; GO:0003824; F:catalytic activity; IEA.

DR GO; GO:0008152; P:metabolism; IEA.

DR InterPro; IPR000873; AMP-bind.

DR Pfam; PF00501; AMP-binding; 1.

DR PROSITE; PS00455; AMP BINDING; 1.

SO SEQUENCE 446 AA; 49317 MW; BAIED75849EDF92B CRC64;

Query Match

Best Local Similarity 54.7%; Score 2103; DB 11; Length 446;

Matches 392; Conservative 26; Mismatches 28; Indels 0; Gaps 0;

QY 285 MGLHMAAGPETHPAGISDLIAEVSABVDGVPVGYLSSPGISDTCTIYFTSGTTGLPRA 344

DB 1 MGLHMAAGPETHPAGISDLIAEVSABVDGVPVGYLSSPGISDTCTIYFTSGTTGLPRA 60

QY 345 ARISHLKLLOCGFYOLCGVHOEDVYIALPLVHMSGSLGIYVCMGIGATVVLKSFSA 404

DB 61 ARISHLKLLOCGFYOLCGVHOEDVYIALPLVHMSGSLGIYVCMGIGATVVLKSFSA 120

QY 405 GQFMEQCOGHVTVFOYIGELCRYLVNPPSKAERGHKVRILAVSGLRPPTWERFRRG 464

DB 121 SQFMDQCOGHVTVFOYIGELCRYLVNPPSKAERGHKVRILAVSGLRPPTWERFRRG 180

QY 465 PLQVLETYGTEGNVATNTYTGGAAGVGRASWLKYHIFPFSILRYDVTGEPTRDQGHG 524

DB 181 PLQVLETYGTEGNVATNTYTGGAAGVGRASWLKYHIFPFSILRYDVTGEPTRDQGHG 240

```

QY 525 MATSREPEGLVAPVSOQSPFLGYAGBELAQKLDKVERPQGVFNTGDLVDCDGF 584
DB 241 MTTSREPEGLVAPVSOQSPFLGYAGBELAQKLDKVERPQGVFNTGDLVDCDGF 300
QY 585 LRFHRTGDTFRMKGNVATTEVAEVEFALDFOENVYGVTPGHEGRAGMAALVLRP 644
DB 301 LRFHRTGDTFRMKGNVATTEVAEVEFALDFOENVYGVTPGHEGRAGMAALVLRP 360
QY 645 HALDLMLQYTHVSENPYPARPRFLRQESLATTTFQOQKVRANEGFPDSTLSDPLV 704
DB 361 QALNTVQLVSHVSENPYPARPRFLRQESLATTTFQOQKVRANEGFPDSTLSDPLV 420
QY 705 LDOAVGAYPLFTFARYSALLAGNLRI 730
DB 421 LDOIGAYPLFTFARYSALLAGNLRI 446

RESULT 5
Q9BTJ5 PRELIMINARY; PRT; 402 AA.
ID Q9BTJ5;
AC Q9BTJ5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE Similar to hypothetical protein WC4365 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003654; AA003654.1;
DR GO; GO:0003824; F:enzymatic activity; IEA.
DR GO; GO:0008152; F:metabolism; IEA.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 1.
DR PROSITE; PS00455; AMP BINDING; 1.
KM Hypothetical protein.
FT NON TER
SQ SEQUENCE 402 AA; 44092 MW; BID946FC9822C3P7 CRC64;

Query Match 54.2%; Score 2083.5; DB 4; Length 402;
Best Local Similarity 91.7%; Pred. No. 4, 4e-135;
Matches 398; Conservative 1; Mismatches 0; Indels 35; Gaps 1;

QY 297 HPAGISDLAEVSAEVDGVPGLTSPQITDTCTCYIFTSQTTGPKAARISHLKIQCQ 356
DB 4 HPAGISDLAEVSAEVDGVPGLTSPQITDTCTCYIFTSQTTGPKAARISHLKIQCQ 63
QY 357 GFYQLCGVHEDVYIALPLYMSSLLGIVGCMGIGATVVKSFSAQGFWECDQCHRV 416
DB 64 GFYQLCGVHEDVYIALPLYMSSLLGIVGCMGIGATVVKSFSAQGFWECDQCHRV 123
QY 417 TYPQYIGELCYLVNPPSKAERGHKVRILAVSGARPDTPWRFVRFGPLQVLETYGLTE 476
DB 124 TYPQYIGELCYLVNPPSKAERGHKVRILAVSGARPDTPWRFVRFGPLQVLETYGLTE 183
QY 477 GNVATINYGORGAVGRASWLYKHIFPESLIRYDVTGTGEPRIIDPOGHCMATSPGEGLLV 536
DB 184 GNVATINYGORGAVGRASWLYKHIFPESLIRYDVTGTGEPRIIDPOGHCMATSPGEGLLV 243
QY 537 APVSQSPFLGYAGBELAQKLDKVERPQGVFNTGDLVDCDGFRLRFHRTGDTFR 596
DB 244 APVSQSPFLGYAGBELAQKLDKVERPQGVFNTGDLVDCDGFRLRFHRTGDTFR 303
QY 597 WKGENVATTEVAEVEFALDFOENVYGVTPGHEGRAGMAALVLRPHALDMLQYTHV 656
DB 304 WKGENVATTEVAEVEFALDFOENVYGVTPGHEGRAGMAALVLRPHALDMLQYTHV 363
QY 657 SENLPPYARPRFLRQESLATTTFQOQKVRANEGFPDSTLSDPLVLDQAVGAYPLT 716

```

```

DB 364 SENLPPYARPRFLRQESLATTTFQOQKVRANEGFPDSTLSDPLVLDQAVGAYPLT 388
QY 717 TARYSALLAGNLRI 730
DB 389 TARYSALLAGNLRI 402

RESULT 6
Q8TEJ0 PRELIMINARY; PRT; 288 AA.
ID Q8TEJ0;
AC Q8TEJ0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE FLJ00207 protein (fragment).
GN FLJ00207.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Ukuya H., Takano Y., Nomura N., Kikuno R., Nagase T., Ohara O.;
RT "The nucleotide sequence of a long cDNA clone isolated from human
RT spleen."
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK074134; BAB84960.1;
DR GO; GO:0003824; F:enzymatic activity; IEA.
DR GO; GO:0008152; F:metabolism; IEA.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 2.
DR PROSITE; PS00443; GATASE_TYPE_II; 1.
FT NON TER
SQ SEQUENCE 288 AA; 29865 MW; 09AD767A6497A95C CRC64;

Query Match 36.6%; Score 1407; DB 4; Length 288;
Best Local Similarity 100.0%; Pred. No. 9e-89;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVCORTRAPWKEKSLERBALGFRKSGMFGASGNNQTVPIEAGSNALILPLILL 60
DB 13 MGVCORTRAPWKEKSLERBALGFRKSGMFGASGNNQTVPIEAGSNALILPLILL 72
QY 61 PLILKILHMPQILRMIPADIAFVRLCCRAIRARALAAAAADPGPGGSLARLAE 120
DB 73 PLILKILHMPQILRMIPADIAFVRLCCRAIRARALAAAAADPGPGGSLARLAE 132
QY 121 LAOORAAHTFLHSGRFRFSYSEARERSNRAARAFLEALGMDWGPDCGDSGSGAGEGERA 180
DB 133 LAOORAAHTFLHSGRFRFSYSEARERSNRAARAFLEALGMDWGPDCGDSGSGAGEGERA 192
QY 181 APAGDAAGSGAEFAGGDAARAGGAAPLSGATVALLLPAGPEFLWLFGLAKGIR 240
DB 193 APAGDAAGSGAEFAGGDAARAGGAAPLSGATVALLLPAGPEFLWLFGLAKGIR 252
QY 241 TAFVPTALRRGRLHCLRSAGARLVLP 269
DB 253 TAFVPTALRRGRLHCLRSAGARLVLP 281

RESULT 7
Q7Z6E6 PRELIMINARY; PRT; 619 AA.
ID Q7Z6E6;
AC Q7Z6E6;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE Hypothetical protein DKFZp779M0564.
GN DKFZp779M0564.
OS Homo sapiens (Human).

```


RESULT 11

OBAYCS PRELIMINARY; PRT; 625 AA.
AC OBAYCS;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Similar to solute carrier family 27 (Fatty acid transporter), member 2 (Fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxId=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strassberg R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC041746; AAH41746.1; -
DR GO; GO:0003824; P:metabolic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR00873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 2.
DR PROSITE; PS00455; AMP_BINDING; 1.
FT NON TER 1
SQ SEQUENCE 625 AA; 71253 MW; E5C22101B2DE51B4 CRC64;

Query Match 34.2%; Score 1312.5; DB 13; Length 625;
Best Local Similarity 39.5%; Pred. No. 7.8e-82;
Matches 268; Conservative 108; Mismatches 234; Indels 69; Gaps 6;
QY 54 LPPLLLPLLLKL--HLPOLRMLPADLAFVRLCKCKRLRAALAAAADDEGEGG 111
DB 14 LPGLLLILLISFIFPIYFQDIAVITAVRGIIR---RVSSTPHTVVD----- 62
QY 112 CSLARLAEALAOQRAHFFLIHGSRRFSISEMERSSNPAAPFLALGMDWGPDGDSGE 171
DB 63 ----MFLKVERHDKPFVLFEEVYVYSHNDKLSNQAARA-DR----- 101
QY 172 GSAGEGEPAAPGAPAAAGSGAEFAGDGAARCGAAPLPGATVALLPAGSEFLTW 231
DB 102 -----KRAATKSGDC-----VAIFMAAARAYITW 126
QY 232 PGLAVALGRTAFVPTALRRGPIILHCLSCGAPALVLAPELESLEPDIPALRANGHLMA 291
DB 127 LGVAALGSGIACLNINIRSSQSLHCFRCRAKVIILAEPELKDVIIEVMPBLRKXNVKVF 186
QY 232 AGPGHHPAGISDLAEVASAVDGPVGLSSPQITDCLYIFPSGQTGLPKAARISLTK 351
DB 187 LTKIVISSETESFLDKVRAADSEVPKSLRFSVSKSLAMTYIVSGITGLPKAALINRYR 246
QY 352 ILCCGFFYQLCGVHEDVITYALPLYHMSGSLIGVCGMIGATVLSKFSASQFWDG 411
DB 247 ILSACGMFEICKVARADVYSPLPLYHSSAMMIGHGICISKAGATVLPKFSASQFWDG 306
QY 412 QOHRITVQYIGELCRILYNOPPSKAERGHKRLAVSGILRPDPTERVVRPFGPIQVLET 471
DB 307 RRYNTVILYIGEIVRLCNVPRKSDDVAAHNVRAIIGGLRTDVSEFLRFGELHNEF 366
QY 472 YGLTGNVATINVTQORGAVGASWLYKHIFPFSILRYDVTGTEIRDPQCHCATSPGE 531
DB 367 VASTGNTIAFINNTVSGVGRFDKILHSYDIKTDIEMDEVRAMRCKAKARQG 426
QY 532 PELLVAPVVSQSPPLGYAGGPELAQGLLKXDFRPGDVFNVTGDLIVDDGGLRPHDXT 591
DB 427 PELLCKINSMPFGVAGDEHSTERKIMRDFRKGDAFYNSGDLITVDQONFIYFHDV 486
QY 592 GPTFRMKGENVATTEVAEVEALDPLQEVNVYGVVPGHEGAGAAALVLRPHALDLMQ 651
DB 487 GPTFRMKGENVATTEVADILCTIVNIOEVNIVGVSVQNHGRIKRAAILLDEEVEFGDK 546
QY 652 LVTHVSENLPPYAPRFLRQESLATTETFKQKVRMANEGDPSTLSDPLVILDQAVGA 711

DB 547 LVNAVDFLNPVAPRPFIRQNSMDITGTFKQKVLVGEFPAISDPLVILDBREKK 606
QY 712 YLPPLTARRISALLAGNLRI 730
DB 607 YPMTQTIVEDIKMKRIKL 625

RESULT 12

Q9Y2P5 PRELIMINARY; PRT; 690 AA.
ID Q9Y2P5
AC Q9Y2P5;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Very long-chain acyl-CoA synthetase homolog 2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99410693; Pubmed=10479480;
RA Steinberg S.U., Wang S.U., McGuinness M.C., Watkins P.A.;
RT "Human liver-specific, very-long-chain acyl-coenzyme A synthetase: cDNA cloning and characterization of a second enzymatically active protein."
RT Mol. Genet. Metab. 68:32-42(1999).
RL EMBL; AF064255; AAD29444.1; -
DR Genew; HGNC:10399; SLC27A5.
DR GO; GO:0005783; C:endoplasmic reticulum; TAS.
DR GO; GO:0004467; F:long-chain-fatty-acid-CoA ligase activity; TAS.
DR GO; GO:0000038; P:very-long-chain fatty acid metabolism; TAS.
DR InterPro; IPR00873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 1.
DR PROSITE; PS00455; AMP_BINDING; 1.
SQ SEQUENCE 690 AA; 75384 MW; 011313424D794546 CRC64;

Query Match 33.9%; Score 1303.5; DB 4; Length 690;
Best Local Similarity 42.1%; Pred. No. 3.7e-81;
Matches 297; Conservative 93; Mismatches 251; Indels 65; Gaps 13;
QY 28 GSGMFPASGM--NQVVP--IEAGSMALLLLPLLLPLLLPLLLKHLMPOLRMLPADLAFV 84
DB 47 GLAMLARPMGPPVPPGLSLAAALALTLP-----ARLPGRMLPADVIFLA 95
QY 85 RALCKRALRARAALAAAADPEGPEGCSLAWRLAEAOQRAHTEFLHGSRRFSYSAE 144
DB 96 KILHLGUKR-----GC-----LSRQPDPTVDAFERR-----AR 125
QY 145 RESNRARARALRLALGMDWGDGDSGEGSAGESEERAPAGADAAAGSGAEFAGDGAAG 204
DB 126 AQGGERA-----LVVT-GP-----GAGSVTFEILDA-RACQAAWLKLEL--GDAASLC 170
QY 205 GGAAPLSPGATVALLPAGPEFLTMFGIAXAGLRTAFVPTALRRGPIILHCLSCGARA 264
DB 171 AG-----EPTALLVLSQAVRA-LCMILGIAGCTMTNINPGRMPLAHSVLSSGARV 224
QY 265 LVLAPEFLESLEPDLPALRMAGHLMAAGGTHPAGISDLAEVASAEVNDGPVPGYISSPQ 324
DB 225 LVVDPDLRESLEIILPKLQENIRCFYLSHTSPGVGALGALDAPSHVPADIRAGI 284
QY 325 SIIDTCYITSGTGTGAPKARISHLKLIIQCGSFYQLCGVHEDVITYALPLYHMSGSL 384
DB 285 TWSPALFITTSGETTGLPKAIIHERVILQMSKSLSGATADVVYTYLPLYHWGLVY 344
QY 385 GIVCGMIGATVVLKSKFSAGQFWDGQOGRVTVQYIGELCRILYNOPPSKAERGHYR 444
DB 345 GILGIDLAGTAVLAPRFSYSCFWDQRGVTVILYVGSILRYLCNIPQOPEDRTHYR 404
QY 445 LAVGSGLRPDTFRFRFRPGLQVLETYGTENAVNTIVTQORGAVGASWLYKHIFP 504
DB 405 LAMNGLRADVWETFPQRFGRIRIWEVYGSTBSMGLVINVVGCGALGKVSCLRLVLSPE 464

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCB1_TaxId=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Salivary gland;
 RA Strausberg R.;
 RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC013335; AAH13335.1; -
 DR EMBL; BC013272; AAH13272.1; -
 DR PIR; J01017; J01017.
 DR MGD; MGI:1347100; S1C27A5.
 DR GO; GO:0003824; F: catalytic activity; IEA.
 DR GO; GO:0008152; P: metabolism; IEA.
 DR InterPro; IPR000873; AMP-bind.
 DR Pfam; PF00501; AMP-binding; 1.
 DR PROSITE; PS00455; AMP_BINDING; 1.
 DR Hypothetical protein.
 SQ SEQUENCE 689 AA; 76202 MW; 1642BBC2CF04FAA3 CRC64;

Query Match 32.6%; Score 1251; DB 11; Length 689;
 Best Local Similarity 39.8%; Pred. No. 1.5e-77;
 Matches 280; Conservative 91; Mismatches 255; Indels 78; Gaps 8;

QY 46 GSMALLPLPLLL-----LPILLKLHLPOLRWLPADLAFARA 86
 DB 37 GDFPLVLGLLALGRPWISSMPPHSLVGAALTLFLPLQPPGLRLMKDVAFTFM 96
 QY 87 LCCRAALPAPALAAAADDEGEGCSLAWELAEIAQORAHTFLHGSRRFSYSEARE 146
 DB 97 LFYGLKFFRR-----LNKHPPE-----TIV-----DALEKQ 122
 QY 147 SNRABAFRLALGMWGDG-----GDSGSGSAGEGERAPAGDAAGSAGAFAGCGAA 202
 DB 123 -----ALAW-----PDRVALVCTGSESSITNSQDARSQDAWYTKKLDVAYIQN 169
 QY 203 RGGGAAPLSPGATVALLPAGPEFLMWFLGAKAGLRTAFTALRRGPLLHCRGCA 262
 DB 170 TRDAALITVLPSTKISALS-----VFLGAKLGCVPAMINPHSRGMFLHSVRSQA 221
 QY 263 RALVIAPEFLSEBDELPAIRAMGHLMAAGPCHPAGISDLAEVSAEYVGPVGLSS 322
 DB 222 SVLIADPDLQENLEEVLPKLAENHCHYLGHSSTPGEVALGSLDAPSDPVPASIRA 281
 QY 323 POSITDCLYFTSGTGLPKAARISHLKIQCQGFYOLCGVHQBVDVYIALPYHMSGS 382
 DB 282 TIKKMSPAIFFTSGTGLPKPALISHRVIOVSNVLSFGCGRADVDVYVLPYHTIGL 341
 QY 383 LLIIVGCMGIGATVVLKSKFSAGQWEDCCQOHRVTVFOYIGELCRYLVNPPSALRGGK 442
 DB 342 VLGFGICQVATCTVLAKFASRMAECROHGVVILVGEILRYLCNVEQPEDKHT 401
 QY 443 VRLAVSGSLRPDIWERFVRFGPLOVETYGTEGNVATINYTGORGAVGRASLYHIF 502
 DB 402 VRLANGNGLRANWKNFOQRFGPIRINEFYGTBGNVGLMYVGHGCAVGTSCILMLT 461
 QY 503 PPSLIHYVTTGEPIRDPQGHCMATSPGEGILVAVPSQSGPFLGYAGPELQAGKLLKD 562
 DB 462 PPELVQDFIETAEPLRDKQGFCEIPEPEKPGILLTKYKNOPLGYSGSQAESNRKLVAN 521
 QY 563 VERPDVFEFNTGDLVCDQDGLRFHNDTGDTPFKMGENVATTEVAVFEALDLQGVNY 622
 DB 581 VRRGDLVFNNGDVLTLDOEGFFYQDLGDTFRKMGENVSTGEVECVLSLDLEEVN 581
 QY 623 YGVTVPGHGRAGMAALVLRPPHALDMQLYTHVSENLPYARPRFLRQSLATTEFTK 682
 DB 582 YGVVPGECEGKVMAAVLAPOKTPDQCKLVHVRSMPLPAYATPHFRIDQSLITNTYK 641
 QY 683 QOKRMAHGFDPSTLSPVLYLDOAVAYPLPTARSAALLAG 726
 DB 642 LVKSLVREGEFVDVGIADPLYLIDNKAQTFRSLMPDVYQAVCEG 685

RESULT 15
 Q95338 PRELIMINARY; PRT; 690 AA.
 ID Q95338
 AC Q95338
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Bile acid CoA ligase.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCB1_TaxId=10116;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley;
 RA Falany C.N., Xie X., Wheeler J., Wang J., Barnes S.;
 RT "Molecular cloning and expression of rat liver bile acid CoA ligase";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF242189; AAG09770.1; -
 DR GO; GO:0016874; F: ligase activity; IEA.
 DR GO; GO:0008152; P: metabolism; IEA.
 DR InterPro; IPR000873; AMP-bind.
 DR Pfam; PF00501; AMP-binding; 1.
 DR PROSITE; PS00455; AMP_BINDING; 1.
 DR Ligase.
 SQ SEQUENCE 690 AA; 76265 MW; D5EBB8317758DA59 CRC64;

Query Match 32.0%; Score 1231; DB 11; Length 690;
 Best Local Similarity 39.2%; Pred. No. 3.6e-76;
 Matches 280; Conservative 98; Mismatches 267; Indels 70; Gaps 12;

QY 52 LLLPLLLPLLLKLHLP-----QLRWLPAD-----LAFVRLCK-----RAL 93
 DB 10 LLLSLILLVG--LGLPLPAPATALLRNFLGDPCTFVLGLAFLGRPWISSITPWL 66
 QY 94 RARALAAAADDEGEGCSLAWRLAEIAQORAHTFLHGSRRFSYSEARESNR-AAR 152
 DB 67 AAALITSLPPRPP-----ELRWLHSDVA-----FAFLILEYG-----LNLRRRLNHPPE 114
 QY 153 APLRALGMWDPDGGSGSGSAGEGERAPAGDAAA-----GSGAFAGCGAARCGGA 207
 DB 115 LFLDAL-----EQOARPDVALVCTGSEGSITNRELNARACQ 155
 QY 208 AAPL-----SPGATVALLLPA-GPEFLMWFLGAKAGLRTAFTALRRGPLH 255
 DB 156 AMLKAKLKEATIQEDKATAILVPSKISALSVLGLAKLGCVPAMINPHSRGMPLH 215
 QY 256 CLRSQARALVLAPEFLSEBDELPAIRAMGHLMAAGPCHPAGISDLAEVSAEVDGP 315
 DB 216 SVQSSASVLIADPDLQENLEEVLPKLAENHCHYLGHSSTPGEVALGSLDAPSDP 275
 QY 316 VPGYLSPOSITFTCTYFTSGTGLPKAARISHLKIQCQGFYOLCGVHQBVDVYIALP 375
 DB 276 VPKLAANKKMSPAIFFTSGTGLPKPALISHRVIOVSNVLSFGCGRADVDVYVLP 335
 QY 376 LYHMSGSLIIVGCMGIGATVVLKSKFSAGQWEDCCQOHRVTVFOYIGELCRYLVNPPS 435
 DB 336 LYHSMGLVIGVGLCTQVLAFCVLAPEFASRMAECROHGVVILVGEILRYLCNVEQ 395
 QY 436 KAERGHKVLAVSGSLRPDIWERFVRFGPLOVETYGTEGNVATINYTGORGAVGRAS 495
 DB 396 PEDKHTVRLANGNGLRANWKNFOQRFGPIRINEFYGTBGNVGLMYVGHGCAVGTSC 455
 QY 496 WLYKHLPPSLIRYVTTGEPIRDPQGHCMATSPGEGILVAVPSQSGPFLGYAGPEL 555
 DB 456 CFIRMLTPELVQDFIETAEPLRDKQGFCEIPEVETGPGILLTKIRKNOPLGYSGQDET 515
 QY 556 QGKLLKDVFRPDVFNNTGDLVCDQDGLRFHNDTGDTPFKMGENVATTEVAVFEALD 615
 DB 516 KRLVANVQVDDLYNNTADVLLADQEGFFYRDRGDTFRKMGENVSTREVEGVLSILD 575
 QY 616 FLQEVNVVTVPGHGRAGMAALVLRPPHALDMQLYTHVSENLPYARPRFLRQSL 675

Db 576 FLEEVNYGVTVPGCEGKVMGMAAVKLAPEGKTFDGGKLYOHVRSWLPAYATPHFIRIODSL 635

Qy 676 ATTEFKOQKVRMANEGFDPSTLSDPYVLDQAVGAVLPLTTARYSALLAGNLR 730

Db 636 EITNTYKLVKSQIAREGFDVGVIADPLYIIONKAKETFRSLMPDVYQAVCEGTWKL 690

Search completed: June 30, 2004, 17:34:58
Job time : 54 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2004, 17:29:24 ; Search time 17 Seconds
(without alignments)
2235.954 Million cell updates/sec

Title: US-10-030-226-2

Percent score: 3843
Sequence: 1 MGVCQRTAPWKEKSQLERA.....AYLPLTARYSALLAGNLRI 720

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1351	35.2	620	VLCS_MOUSE	O35488 mus musculus
2	1340	34.9	620	VLCS_RAT	P97524 rattus norv
3	1339.5	34.9	620	VLCS_HUMAN	O14975 homo sapien
4	1056.5	27.5	646	PATP_MOUSE	O60714 mus musculu
5	1052.5	27.4	646	PATP_RAT	P97849 rattus norv
6	654	17.0	623	FAT1_YEAST	P98225 saccharomyc
7	405.5	10.6	522	CAIC_ECO57	Q8xa54 escherichia
8	393.5	10.2	522	CAIC_ECOLI	P1552 escherichia
9	353.5	9.2	517	CAIC_SALTI	Q82914 salmonella
10	352.5	9.2	517	CAIC_SALTY	O82914 salmonella
11	276.5	7.2	561	LCFA_YERPE	O82914 salmonella
12	273.5	7.0	561	LCFA_ECO57	Q8xa54 escherichia
13	269.5	6.9	561	LCFA_ECOLI	P29212 escherichia
14	265.5	6.9	561	LCFA_SALTY	Q82914 salmonella
15	259	6.7	485	MENE_ENTFA	Q83881 enterococcu
16	254.5	6.6	548	YDID_ECOLI	P38135 escherichia
17	251	6.5	1088	TYCA_BRSPA	P09095 brevbacillu
18	247	6.4	543	PAT2_YEAST	P38137 saccharomyc
19	246.5	6.4	3587	TYCB_BRSPA	O30408 b tyrocidin
20	242	6.3	2560	PP82_BACSU	P29846 bacillus su
21	239.5	6.2	6359	BACQ_BACLI	O68008 b bacillraci
22	239	6.2	5255	BACA_BACLI	O68008 b bacillraci
23	237.5	6.2	563	ACLI_ORYSA	P17814 oryza sativ
24	233.5	6.1	1293	ENTF_ECO57	O82914 salmonella
25	233	6.1	537	ACL_PINTA	P16166 pinus taeda
26	231	6.0	572	ACSA_BACSU	P39052 bacillus su
27	230.5	6.0	1281	ENTF_SHIFL	P29868 shigella fl
28	230.5	6.0	1293	ENTF_ECOLI	P11454 escherichia
29	225	5.9	6486	TYCC_BRSPA	O30409 b tyrocidin
30	222.5	5.8	672	ACSA_PHYBL	Q01576 phycomyces
31	222	5.8	486	MENE_BACSU	P33971 bacillus su
32	220	5.7	545	ACLI2_SOLITU	P1685 solanum tub
33	220	5.7	547	ACLI1_TOBAC	O24145 nicotiana t

34	219.5	5.7	651	1	ACSA_MYCTU	O69635 mycobacteri
35	219.5	5.7	651	1	ACSA_STRCO	Q9x928 streptomyc
36	219	5.7	545	1	ACLI2_SOLITU	P31684 solanum tub
37	219	5.7	649	1	ACSA_DEIRA	O9x117 deinococcus
38	218.5	5.7	474	1	MENE_STREP	O82821 streptococ
39	218.5	5.7	562	1	LCFA_HAEIN	P46450 haemophilus
40	218.5	5.7	651	1	ACSA_MYCBO	P59871 mycobacteri
41	217.5	5.7	661	1	ACSA_COPCI	O13440 coprinus ci
42	217.5	5.7	3588	1	SREP1_BACSU	P27206 bacillus su
43	216.5	5.6	560	1	LCFA_BACSU	P94547 bacillus su
44	216	5.6	656	1	ACSA_RHOCA	O68040 rhodobacter
45	214	5.6	645	1	ACSA_PSEAE	Q9hV66 pseudomonas

ALIGNMENTS

RESULT 1	ID	VLCS_MOUSE	STANDARD;	PRT;	620 AA.
AC	O35488; 070550;				
DT	15-JUL-1998 (Rel. 36, Created)				
DT	15-JUL-1998 (Rel. 36, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Very-long-chain acyl-CoA synthetase (EC 6.2.1.-) (Very-long-chain-fatty-acid-CoA ligase).				
GN	SLC27A2 OR FAVCL1 OR VLCS OR VLCS.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=BALB/c; TISSUE=Liver;				
RA	Kemp S., Lu J.-F., Smith K.D.;				
RU	Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=BALB/c; TISSUE=Liver;				
RX	MEDLINE=98218572; PubMed=9559670;				
RA	Berger J., Truppe C., Neumann H., Foxes-Petter S.;				
RT	"cdna cloning and mRNA distribution of a mouse very long-chain acyl-CoA synthetase.";				
RL	FEBS Lett. 425:305-309(1998).				
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein. Peroxisomal.				
CC	-1- TISSUE SPECIFICITY: Strong expression in liver and kidney, low expression in brain and testis, no expression in skeletal muscle and spleen.				
CC	-1- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme family.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).				
CC	EMBL: AF033031; AAB87982.1; -				
DR	EMBL: AJ223958; CA11687.1; -				
DR	MGD: MG:1347099; S1C27A2.				
DR	InterPro: IPR000873; AMP-bind.				
DR	Pfam: PF00501; AMP-binding; 1.				
DR	PROSITE: PS00455; AMP BINDING; 1.				
KW	Ligase; Fatty acid metabolism; Peroxisome; Transmembrane.				
FT	TRANSMEM 1 21				
FT	POTENTIAL.				
FT	TRANSMEM 107 127				
FT	POTENTIAL.				
FT	TRANSMEM 262 282				
FT	POTENTIAL.				
FT	CONFICT 35 35 Q -> R (IN REF. 2).				
SQ	SEQUENCE 620 AA; 70366 MW; 77C9BBD0DB39FFB CRC64;				
Query Match	35.2%; Score 1351; DB 1; Length 620;				


```

Db 250 SGTGLPAAIVSVSRYYRIAFHHSYSKMANVLDCLFLYHSANINGVGCCITTYGL 309
Qy 395 TVVLKSKFSAGQFWECCOHRVTVFOYIGELCRYLVNVPSPKABRGHKVRLAVSSGLRPD 454
Db 310 TVLVRKFSASRPFMDCKVNCYVQYIGELCRYLVNVPSPKABRGHKVRLAVSSGLRPD 369
Qy 455 TWSPFRARFGLQVLEYIGLTVGNVATNTYQGRGANGRSMVLYKHIFPSLITVDYTTG 514
Db 370 TWSPFRARFGLQVLEYIGLTVGNVATNTYQGRGANGRSMVLYKHIFPSLITVDYTTG 429
Qy 515 EPIDPOGHCMATSPGEGGLVAPVQSSP---FLGAVAGBELAQGLKLDVFPFGVVF 571
Db 430 EPLDSSQCLICPCQGEFGLVQINQDDPLRFDDGV-SLSANKKILASVFRKGSAY 488
Qy 572 NTGDLVLDGQFLRFHRTGDTFRMKGENVATTEVAEVEBALDPLQEVNVTYVPGHE 631
Db 489 LSGDVLVMDLGYVWFDRSGDTFRMKGENVATTEVAEVEBALDPLQEVNVTYVPGHE 548
Qy 632 GRAGMAALVLRPPHALDMQLYTVHSENLPPYARPRFLRQESLATTETPRQGVRRANE 691
Db 549 GKSGMMA-LADPHQLDPSNMYQELQKVLASVAPIFLRLLPOVDTTGTFKIQKRLQRE 607
Qy 692 GFDPSTLSDPLVYLDQAVGAYLPLTARYSALLAGNLR 730
Db 608 GFDPSTLSDPLVYLDQAVGAYLPLTARYSALLAGNLR 646

```

RESULT 6 FAT1 YEAST STANDARD; PRT; 623 AA.

```

AC P38225;
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 40, Last annotation update)
DE Probable long-chain fatty acid transport protein.
GN FAT1 OR YBR041W OR YBR0411.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
CX NCBI_Taxid=4932;
RN 11
RS SEQUENCE FROM N.A.
RC STRAIN=S288c;
RA Andre B., Czechluch C., Hein C., Jauniaux J.C., Urrestarazu A.,
RA Viessers S.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RP CHARACTERIZATION.
RC STRAIN=W303A;
RX MEDLINE=97236810; PubMed=9079682;
RA Paargeman N.J., Diruseo C.C., Elberger A., Knudsen J., Black P.N.;
RT "Disruption of the Saccharomyces cerevisiae homolog of the murine
RT fatty acid transport protein impairs uptake and growth on long-chain
RT fatty acids.";
RL J. Biol. Chem. 272:8531-8538 (1997).
CC -1- FUNCTION: May be involved in long-chain fatty acids uptake, and
CC thus may play a pivotal role in regulating their accessibility
CC uptake of these hydrophobic compounds under conditions where fatty
CC acid synthesis is compromised.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).

```

```

DR EMBL; Z35910; CAA84983.1; -.
DR PIR; S45899; S45899.
DR GeneBank; U00002.45; FAT1.
DR GO; GO:0005811; C:lipid particle; IDA.
DR GO; GO:0005792; C:micosome; IDA.
DR GO; GO:0005777; C:peroxisome; IDA.
DR GO; GO:0005886; C:plasma membrane; IMP.
DR GO; GO:0005324; F:long-chain fatty acid transporter activity; IMP.
DR GO; GO:0004467; F:long-chain-fatty-acid-CoA ligase activity; IMP.
DR GO; GO:0006869; P:lipid transport; IMP.
DR GO; GO:0000038; P:very-long-chain fatty acid metabolism; IMP.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 1.
DR PROSITE; PS00455; AMP BINDING; 1.
KW Lipid transport; Transmembrane.
FT TRANSMEM 6 26
FT TRANSMEM 54 71
FT TRANSMEM 149 169
FT TRANSMEM 293 313
FT CARBOHYD 184 184
FT CARBOHYD 289 289
FT CARBOHYD 534 534
FT CARBOHYD 591 591
SO SEQUENCE 623 AA; 71697 MW; 0AE0270B60C8CFE CRC64;

```

Query Match 17.0%; Score 654; DB 1; Length 623;
Best Local Similarity 33.7%; Pred. No. 2.9e-36;
Matches 169; Conservative 84; Mismatches 193; Indels 56; Gaps 14;

```

Qy 214 GATVALLPAGPEFLTMFGLAKAGLRTAFVPTALRGFLHCLRSQGARALVLP--- 269
Db 137 GDVAIDCTNKPFLFVFLWLMSTWNIGAPFLFVNNYKGTPLVSLKISNTQYFIDPDASN 196
Qy 270 ---EFIESLEPDLPALRAMGLHMAAGPETHAGISDLIAVSABVDGVPGLS----- 321
Db 197 PIRESESEIKNALPDVKNLVLE-----EQDLMEHE---LNSQSPFELQDDNR 241
Qy 322 SPOSITD--TCLYIFSGTGLPKAARISHLK-LIQCQGFYGLCGHQBVDVYLLPLVH 378
Db 242 TPLGLDFPKSMILYISGTLGPKALMSKRSVGCQVFGVNLMTNISTVTFAMPLFH 301
Qy 379 MGSGLIGVCMGIGATVVLKSKFSAGQFWECCOHRVTVFOYIGELCRYLVNPPSKAE 438
Db 302 STALLGACALISHGCLASHKFSASTFWKQVYLGAHIGYVSEVCRYLHTPTISKYE 361
Qy 439 RGHVYFLAVSGLRPDPTRFVRFRGGLQVLEYIGLTVGNVATNTYQGRGANGRSMV 495
Db 362 KQHKVAVAGNGRPDIWDFRFRNIEVIGFPAATEAPFAATTF--QKGFPGIGACRN 419
Qy 496 -----WLYKHIFPSLIRYDVTGEP-LRDPQGHMATSPEGGLVLA---PVSQOSP 544
Db 420 YGIIQWFLS--FQGLVAMDNDSDVIRNSKGCCEVAPVGEPEMLMRIFPKKPEIS 477
Qy 545 FLGAGGPELAQGLIKLYFRPGDYFNTGDLVLDGQFLRFHRTGDTFRMKGENVAT 604
Db 478 FQGYLNAKETSKXKVRDVFRRGDAMWYRCGLDKKDEYGLWYFLDMRGDTFRMKGENVAT 537
Qy 605 TEVAEVEALDPLQ---EVNVTYVPGHEGRAGMALVLRPPHALDMQLYTVHSENL 658
Db 538 TEVEDGLTNSKEQVAVVGVGKPKYGRGAFVYKL-TNSLIDITAKTLNLSLSR 596
Qy 659 -NIPYARPRFLRQESLATTTE 679
Db 597 LNDPSYAMPFLFYFVDEIKMTD 618

```

RESULT 7
CAIC_ECO57 STANDARD; PRT; 522 AA.
ID CAIC_ECO57
AC Q9XA34;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)

28-FEB-2003 (rel. 41, last annotation update) (EC 6.3.2.-).
 DE Probable crotonobetaine/carnitine-CoA ligase (EC 6.3.2.-).
 GN CAIC OR 20043 OR ECS0040.
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=83334;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Postel E.J., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimallanta E.T., Potamoustis K.,
 RA Apodaca J., Mantharavan T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.,
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7",
 RL Nature 409:529-533(2001).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952;
 RX MEDLINE=21156331; PubMed=11258796.
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasekawa C., Ogasawara N., Yasunaga T.,
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.,
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
 O157:H7 and genomic comparison with a laboratory strain K-12".
 RL DNA Res. 8:11-22(2001).
 CC -1- FUNCTION: Could catalyze the transfer of CoA to crotonobetaine
 or carnitine (By similarity).
 CC -1- PATHWAY: Carnitine metabolism (conversion of carnitine to
 gamma-butyrobetaine).
 CC -1- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as their content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sb.ch/announce/>
 or send an email to license@isb-sb.ch).
 CC -----
 DR EMBL: AE005180; AAG54340.1; -
 DR EMBL: AP002550; BAB33463.1; -
 DR PIR: H85484; H85484.
 DR PIR: H90633; H90633.
 DR InterPro: IPR000873; AMP-bind.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP_BINDING; 1.
 KW Ligase; Complete proteome.
 SQ SEQUENCE 522 AA; 59144 MW; D0F0AB1488B7D8C9 CRC64;
 Query Match 10.6%; Score 405.5; DB 1; Length 522;
 Best Local Similarity 25.8%; Pred. No. 8,9e-20;
 Matches 154; Conservative 78; Mismatches 235; Indels 131; Gaps 19;
 QY 110 GGCSLAWRLAELAQGAATPLIHGS-----RRFSYSAEESRPARAFRLALGMDWGP 164
 DB 10 GGCHLRQWMDLADYGHKHTALICSSGGVNRYSYELNOEIKRANLFFY-TLG----- 63
 QY 165 DGGSDGSGSAGSGERAAPGADMAAGSAGAFAGGAGARGGGAAPSPGATVALLPAG 224
 DB 64 -----IRKGRKVALHLDNC 77
 QY 225 PEFLLWTFGLAVAGLRTAFVPTALRGLLCLRSRGARLVLA-----PEFLSIEPDLF 280
 DB 78 PEFIFCWGGLAKIGIIMVPIVARLLRESAWLQNSQACLIVTSQGFPMYQOQCOEDAT 137
 QY 281 ALRANGHLHMAAGPGRTHRAGISDLLAEVSAVDGVPFY-----LSSPGSTTD 328

138 QLR-----HI-----CLTDVALPADGVSSTOLKNOQAPATLCYAPLSTDD 179
 QY 329 TCIYITSGTGTGPRARISHUKIIQCGFYOL--CGHODEVYLALPLYMSSGLGI 386
 DB 180 TAEILFTSGTTPRPGVITHTN-LRPAGYSAMOCALRDDVYITWPAFIIDOCCTAA 238
 QY 387 VCGMGIGATVVLKSKESAGQFWECCQHRVYFOYIGELCRYLVNOPSKAERGHVR-L 445
 DB 239 MAFSAGATFVLVEKXSARAFWGOVKYRATITECIPMMIRITLMQPSANDRHLREV 298
 QY 446 AAGSGIRPTWTERFPRRFPLOVLETYGTTEBNATIVYTGQ-----GAVGSASMY 498
 DB 299 MFXNLINSBOEKTFCEFRFG-VLLTSYGTETIVGII--GRPEDKRWPSIGRAGCY 354
 QY 499 KILFPFLIRYDVTGEPFRDPQGHQMATSPGEPG-LVAPVSOQSPFLGVAGGPELAQ 557
 DB 355 -----DAEIRD--DINRPLPAEIEICIKVPGKTIREFYLNPK-ATA 396
 QY 558 KILKDFPRPDVFENTGDLIVDDQGLRPHDRTGDTFRMKEENATTEVAEVPALDPL 617
 DB 397 KYLE-----ADGMLHTGDTGYRDERGFYFIDRCNMIKRGGENVSCVELENIATHPKI 451
 QY 618 QEKNVYGVTVPGHEGRAGMALVLRPHALDMQYTHVSENLPYARPRFLQEST 675
 DB 452 QDIIVVGIR-DSIRDEALRAFVVLNBERLTSEEFRCQEMAKFVPSYIEIKDL 508
 RESULT 8
 CAIC_ECOLI STANDARD; PRT; 522 AA.
 ID ID P11552;
 AC AC P11552;
 DT DT 01-JUL-1993 (Rel. 26, Created)
 DT DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE DE Probable crotonobetaine/carnitine-CoA ligase (EC 6.3.2.-).
 GN GN CAIC OR B0037.
 OS OS Escherichia coli.
 OC OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX OX NCBI_TaxID=562;
 RN RN (1)
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC STRAIN=O4:K74;
 RX MEDLINE=95115548; PubMed=7815937;
 RA Eichler K., Bourgis F., Buchet A., Kleber H.-P.,
 RA Mandrand-Berthelot M.-A.,
 RT "Molecular characterization of the cat operon necessary for carnitine
 metabolism in Escherichia coli.",
 RL Mol. Microbiol. 13:775-786(1994).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=92334977; PubMed=1630901;
 RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,
 RA Isono K., Mizobuchi K., Nakata A.,
 RT "Systematic sequencing of the Escherichia coli genome: analysis of
 the 0-2.4 min region.",
 RL Nucleic Acids Res. 20:3305-3308(1992).
 RN (3)
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=9742617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.,
 RT "The complete genome sequence of Escherichia coli K-12".
 RL Science 277:11453-11474(1997).
 CC -1- FUNCTION: Could catalyze the transfer of CoA to crotonobetaine
 or carnitine.
 CC -1- PATHWAY: Carnitine metabolism (conversion of carnitine to
 gamma-butyrobetaine).
 CC -1- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme

CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X73904; CA52113.1; -
DR EMBL; D10483; BA96606.1; -
DR EMBL; AE000114; AAC73148.1; -
DR PIR; B64724; S40558.
DR HSSP; P08659; 1LC1.
DR Ecogen; BG11558; calC.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 1.
DR PROSITE; PS00455; AMP_BINDING; 1.
KM Ligase; Complete proteome.
FT VARIANT 103 103 C -> R (IN STRAIN O44:K74).
FT VARIANT 103 103 RA -> PR (IN STRAIN O44:K74).
FT VARIANT 257 258 Q -> R (IN STRAIN O44:K74).
FT VARIANT 291 291 Q -> S (IN STRAIN O44:K74).
FT VARIANT 333 333 I -> S (IN STRAIN O44:K74).
FT VARIANT 350 350 V -> A (IN STRAIN O44:K74).
FT VARIANT 379 379 I -> V (IN STRAIN O44:K74).
FT VARIANT 393 393 Q -> K (IN STRAIN O44:K74).
FT VARIANT 401 401 A -> V (IN STRAIN O44:K74).
FT VARIANT 413 413 R -> C (IN STRAIN O44:K74).
FT VARIANT 417 417 D -> G (IN STRAIN O44:K74).
FT VARIANT 447 447 A -> T (IN STRAIN O44:K74).
FT VARIANT 508 508 L -> LDI (IN STRAIN O44:K74).
SQ SEQUENCE 522 AA; 59089 MW; 1C54D003B130671 CRC64;
Query Match 10.2%; Score 393.5; DB 1; Length 522;
Best Local Similarity 25.7%; Pred. No. 5.6e-19;
Matches 155; Conservative 79; Mismatches 227; Indels 143; Gaps 22;
QY 110 GGCSLALALALAOAARHTPLHGS-----REFSYSAESESRAARAFIRALGWMGP 164
DB 10 GGQHLRQWMDLADVYGHKTALICSSGGVNVSYELNIEIRTNLFY-TLG----- 63
QY 165 DCGSGESAGEGERAARAGDAAGAAGAEFAGDGAARGGGAAPSPGATVALLPAG 224
DB 64 -----IRKGDVXALHDNC 77
QY 225 PEFLMLWGLAKAGIRAFVPTALRGGPLHC-----LNSCGARALVLA---PEFLS 274
DB 78 PEFLFCWGLAKIG--ALMVPINR---LLCESAMWLNQSAQLVTSQFYPMYQOI 131
QY 275 LEPLPALRAMGLHMAAGPETHPAGISDLAEVSAEDGVDPVX-----LSS 322
DB 132 QOEBAIQIR---HI-----CLTDVALPADDSVSSFTQXNQOPATLCYAP 173
QY 323 POSTTDTCLVYTSQGTGLPRAARISHUKILOCGFYQL--CGVHOEDVIYALPLYMS 380
DB 174 PLSTDDTDIELIFTSQGTSTRPKGVVITHYN-LRFAGYSAWQCALRDDVYLVTPAHID 232
QY 381 GSLIGIVCGMIGTAVYLKSKFSAGOFMEDCOQRHVYFYQYIGLCRYLVNQPSSXERG 440
DB 233 COCTRANMAAFSAGAFVLEKYSARAFVQVQKRAVTECTPMWIRLWQVPSANDQO 292
QY 441 HKVR-LAVSGLRPDTWERFVRRCPLQVLETVGLTEGNVATINYTCR-----GAVG 492
DB 293 HRLREVMYLYNLSEGEKCAFCEFRG-VALLTSYGTEITVIGII--GDRPEDKRWPSIG 348
QY 493 RASWYIKRIFFPSLRIVYVTGEPFRDPQGCMAVSPREP-LVAVYSQOSPFLGYAGG 551
DB 349 RVGFYE-----AEIRD--DHNRLPAEIEIGIKIGPGTKIFKEFLN 391
QY 552 PELAGKLLKDVFRGDFVFNVTGDLVCDQGFTRFHDRTGTRMKGENVATTEVAEVF 611
DB 392 PG-ATAKYLE-----ADGMLHTGDTGYRDEDFYFYVDRRCNMKRGENVSCVELNII 445

QY 612 EALDFLOEVNVYGVTPGHEGRAGMALVLRPHALDMLQLYHVSENLPYARPREFLR 671
DB 446 AAHPKQDIYVVGIK-DSTRDEAIKAFVYLINSEETLSEEFPRFCQNMAKFRVPSYLSI 504
QY 672 QESL 675
DB 505 RKDL 508
RESULT 9
CALC_SALTI STANDARD; PRT, 517 AA.
ID CALC_SALTI
AC Q829J4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable crotonobetaine/carnitine-CoA ligase (EC 6.3.2.-).
GN CALC OR STY0081 OR T0072.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxId=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Barrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holtroyd S., Tagels K.,
RA Krogan A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M.A., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22513167; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G., Iri, Mayhew G.P., Rose D.J.,
RA Burdand V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18.";
RL J. Bacteriol. 185:2330-2337(2003).
CC -i- FUNCTION: Could catalyze the transfer of CoA to crotonobetaine
CC or carnitine (By similarity).
CC -i- PATHWAY: Carnitine metabolism (conversion of carnitine to
CC gamma-butyrobetaine).
CC -i- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL627265; CAD01225.1; -
DR EMBL; AE016834; AA067805.1; -
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 1.
DR PROSITE; PS00455; AMP_BINDING; 1.
KM Ligase; Complete proteome.
SQ SEQUENCE 517 AA; 58509 MW; 547FC052A8FC1512 CRC64;
Query Match 9.2%; Score 353.5; DB 1; Length 517;
Best Local Similarity 24.4%; Pred. No. 2.6e-16;
Matches 144; Conservative 84; Mismatches 248; Indels 113; Gaps 18;

```

QY 110 GGCSLAMRLAEIAOQRAHTFLIHGS-----RRFSYSEARESNRARAFLALGMDWGP 164
D 5 GGONLRQWMDDLAAGVGDKXALIFESCGIYRQFYSALINEINRTALFY----- 55
QY 165 DGGDSGEGSAGEGERAPAGDPAAGSAGEFAGGGAAPLSPGATVALLPAG 224
D 56 -----YLG-----IRKGRVALHLDNC 72
QY 225 PEFILMPLGALAGLRTAFVPTALRRGFLHCLRGSGARALYLADEFI-----ESLEP 277
D 73 PEFICWFGLAKIGALIMPINARLLGSESAMLQNSQVSLVTSQFPMYREIRQDNST 132
QY 278 DLPALRANGHLMAAGPGTHPAGISDLAEVSAEVDGVPYGLSSPGISITDTCLYIFTSG 337
D 133 PLNHICLIGEQLPADGVSH---FSQLQARQATL-----CYPTALSTDDTAELIFTSG 183
QY 338 TTGLPKARASHLKLQCCGFYQL--CGVHODEVYALPLVYHSGSLIGVCGMGAT 395
D 184 TTSRPGGVITHTN-LRFAGYSAMQIALRDDDVYVWPAFHIDCCTAAMPASAGST 242
QY 396 VVLKSKFSAGQFWECCQHRVTVPQYIGELCRYLVNPPSKAERGHKVR-LAVSGLRPD 454
D 243 FVLEKYSARAFWQVRKYQATVTECIPMTITLWQVATPTDRQHHLREVWFYLNLSAQ 302
QY 455 TWRFVRRPGPLQVETYGTEGNVATINYTOR-----GAVGRASWLYKHIFPFLI 507
D 303 EKDAFTERG-VRLITSYGMTETIVGIT---GDRPDGRKRMPSIGRVGSYE----- 350
QY 508 RYDVTGEBIRDPQCHCATSPGEPG-LIVAPVSOQSPFLGAGPBLAQKLDVFRP 566
D 351 -----AEIRDDQNRPLPA--GEIGELCIKIGIPKTIKXYMOP-ATAKALBP----- 396
QY 567 GGVFNTGDLVCCDQGFIRFHDRTGDTFRMKGENVATTEVAEVEALDFLOEVNYYGT 626
D 397 -BGMVHTGDSGYDEBEGFYFVDRRCNMIXRGENVSCVELENIISAHPKIDIVVGIK 455
QY 627 VPGHEGRAGMALVLRPHALDMQLYTVHSENLPYARPRFLRQESL 675
D 456 -DAIRDEAIKAFIVLNEGHTLSEAFPSFCENNAKFKVPSMEIRIDL 503

RESULT 10
CAIC_SALTY STANDARD; PRT; 517 AA.
AC 08ZEX4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Probable crotonobetaine/carnitine-CoA ligase (EC 6.3.2.-).
GN CAIC OR STM0071.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Jorvolik S., Ali J., Dante M., Du F., Hou S., Jayman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2";
RL Nature 413:852-856 (2001).
CC -I- FUNCTION: Could catalyze the transfer of CoA to crotonobetaine
CC or carnitine (By similarity).
CC -I- PATHWAY: Carnitine metabolism (conversion of carnitine to
CC gamma-butyrobetaine).
CC -I- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
CC family.

```

```

CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE008696; AAL1903.1; -.
DR StvGene; SG72727; caic.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 1.
DR PROSITE; PS00455; AMP-BINDING; 1.
KW Ligase; Complete proteome.
SQ SSOURCE 517 AA; 58487 MW; 7F980E9891C67634 CRC64;

Query Match 9.2%; Score 352.5; DB 1; Length 517;
Best Local Similarity 24.6%; Pred. No. 3e-16;
Matches 145; Conservative 84; Mismatches 247; Indels 113; Gaps 18;

QY 110 GGCSLAMRLAEIAOQRAHTFLIHGS-----RRFSYSEARESNRARAFLALGMDWGP 164
D 5 GGONLRQWMDDLAAGVGDKXALIFESCGIYRQFYSALINEINRTALFY-SLG----- 58
QY 165 DGGDSGEGSAGEGERAPAGDPAAGSAGEFAGGGAAPLSPGATVALLPAG 224
D 59 -----IRKGRVALHLDNC 72
QY 225 PEFILMPLGALAGLRTAFVPTALRRGFLHCLRGSGARALYLADEFI-----ESLEP 277
D 73 PEFICWFGLAKIGALIMPINARLLGSESAMLQNSQVSLVTSQFPMYREIRQDNST 132
QY 278 DLPALRANGHLMAAGPGTHPAGISDLAEVSAEVDGVPYGLSSPGISITDTCLYIFTSG 337
D 133 PLNHICLIGEQLPADGVSH---FTOLQARQATL-----CYPTALSTDDTAELIFTSG 183
QY 338 TTGLPKARASHLKLQCCGFYQL--CGVHODEVYALPLVYHSGSLIGVCGMGAT 395
D 184 TTSRPGGVITHTN-LRFAGYSAMQIALRDDDVYVWPAFHIDCCTAAMPASAGST 242
QY 396 VVLKSKFSAGQFWECCQHRVTVPQYIGELCRYLVNPPSKAERGHKVR-LAVSGLRPD 454
D 243 FVLEKYSARAFWQVRKYQATVTECIPMTITLWQVATPTDRQHHLREVWFYLNLSAQ 302
QY 455 TWRFVRRPGPLQVETYGTEGNVATINYTOR-----GAVGRASWLYKHIFPFLI 507
D 303 EKDAFTERG-VRLITSYGMTETIVGIT---GDRPDGRKRMPSIGRVGSYE----- 350
QY 508 RYDVTGEBIRDPQCHCATSPGEPG-LIVAPVSOQSPFLGAGPBLAQKLDVFRP 566
D 351 -----AEIRDDQNRPLPA--GEIGELCIKIGIPKTIKXYMOP-ATAKALBP----- 396
QY 567 GGVFNTGDLVCCDQGFIRFHDRTGDTFRMKGENVATTEVAEVEALDFLOEVNYYGT 626
D 397 -BGMVHTGDSGYDEBEGFYFVDRRCNMIXRGENVSCVELENIISAHPKIDIVVGIK 455
QY 627 VPGHEGRAGMALVLRPHALDMQLYTVHSENLPYARPRFLRQESL 675
D 456 -DAIRDEAIKAFIVLNEGHTLSEAFPSFCENNAKFKVPSMEIRIDL 503

RESULT 11
LCFA_YERPE STANDARD; PRT; 562 AA.
AC 08ZES9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3) (Long-chain acyl-CoA
DE synthetase).
GN FADD OR YPO2074 OR Y2236.
OS Versinia pectis.

```

OC	Bacteria: Proteobacteria: Gammaproteobacteria: Enterobacteriales:
OC	Enterobacteriaceae: Yersinia.
OC	NCBI_taxid=632;
RP	(1)
RP	SEQUENCE FROM N.A.
RP	STRAIN=CO-92 / Biovar Orientalis;
RC	MEDLINE=21470413; PubMed=11586360;
RA	Parthill J., Wren B.W., Thomson N.R., Tibball R.W., Holden M.T.G.,
RA	Baker S., Basham D., Bentley S.D., Brooks K., Churcher C., Murgall K.L.,
RA	Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA	Feltham T., Hamlin N., Holtroyd S., Jagsels K., Kariyev A.V.,
RA	Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.,
RA	Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.,
RT	"Genome sequence of Yersinia pestis, the causative agent of plague.",
RT	Nature 413:523-527(2001).
RN	(2)
RN	SEQUENCE FROM N.A.
RP	STRAIN=KIMS / Biovar Mediaevalis;
RC	MEDLINE=22137863; PubMed=12142430;
RA	Deng W., Burland V., Plunkett G., III, Boutin A., Mayhew G.F., Liss P.,
RA	Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA	Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA	Straley S.C., McDonough K.A., Nilles M.L., Watson J.S., Blattner F.R.,
RA	Perry R.D.;
RT	"Genome sequence of Yersinia pestis KIM.",
RL	J. Bacteriol. 184:4601-4611(2002).
CC	-1- FUNCTION: Esterification, concomitant with transport, of exogenous
CC	for subsequent degradation or incorporation into phospholipids (By
CC	similarity).
CC	-1- CATALYTIC ACTIVITY: ATP + a long-chain carboxylic acid + CoA = AMP
CC	+ diphosphate + an acyl-CoA.
CC	-1- COPACITOR: Magnesium (By similarity).
CC	-1- SUBUNIT: Homodimer (Probable).
CC	-1- SUBCELLULAR LOCATION: Partially membrane-associated (Potential).
CC	-1- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
CC	family.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (see http://www.isb-sdb.ch/announce/
CC	or send an email to license@isb-sdb.ch).
CC	-----
DR	EMBL; AJ14151; CAC9086.1; -
DR	EMBL; AE013826; AAM85796.1; ALT_INIT.
DR	PIR; AB0253; AB0253.
DR	InterPro; IPR000873; AMP-bind.
DR	Pfam; PF00501; AMP-binding; 1.
DR	PROSITE; PS00455; AMP BINDING; 1.
KW	Ligase; Fatty acid metabolism; Magnesium; ATP-binding; Membrane;
KW	Complete proteome.
NP	NP BIND 213 224 ATP (PROBABLE).
FT	VARIANT 504 504 N->K (IN STRAIN KIMS).
FT	SEQUENCE 562 AA: 62639 MW; 812A872B3713D3FD CRC64;
QY	Query Match 7.2%; Score 276.5; DB 1; Length 562;
QY	Best Local Similarity 24.8%; Pred. 3.9e-11;
QY	Matches 131; Conservative 69; Mismatches 234; Indels 95; Gaps 18;
DB	211 LSPGATVALLPAGBEFLMFLGAKAGRTAFVFTALRQGLHCLRSCGARAIVLAP 270
DB	71 LQKSDRALVAMNMLLQYDIAFGVLRACMIYVNNPILTPRELEQLSDSGAVIAIVSN 130
QY	271 FLESLEPDL-----PALRANGVILMAAGPQTHAGISDLAEASAEVDGVPQYSSP 323
DB	131 FAHLEKVFETQVQHVILITMGDL-SAAKGT-----LVNPFVKYIKXIVPYY-LP 181
QY	324 QSI-----TDICLYIFTSQTGLPKAARISH---IKILQCCGF 358

Db	182	DAISPRVTLQGRGRNQYVKPDVIVNTDPAFLQYTGSTTGVANGCALILTHRNMSNLQAKRA	241
Qy	359	YOLGCVHVEDVIYIALPLPYHNSGSLIGVGMGI---	GATVILKSKFSAGQFEDCCQH 414
Db	242	YAPLLQPPRRDILVVALPLPYHIFAL---	TVNCILFETELQGRSLILTNPRDIPGMKELSR 298
Qy	415	RVTVEQYVIGCELCRYLVNPPFKARSGHKVRLAVSSGLR----	PDWERFVRFRGPIQVL 469
Db	229	PFTALGVNTLFPNALINNEEFTHLDFSLRLSVGGMPVQAAVAKETLTGK----	HL 354
Qy	470	ETYGITGEGNVATINYTGQRGAVGRASWLYKHIFPFSILRYDVTGGEPI-----	RDPOG 522
Db	355	EGYGLTECSPLV---TGN-----	PYDLKHVSGSIGLPVSTDRVLRDDG 396
Qy	523	HOMATSPCEPGLVAPVPSQGSPELGYAGPELAGSKLKKVFRGVDVFENMGDLVLCDDQ	582
Db	397	NDEL--GKPELTM--VRGPVMTLGYMORPATD----	DVLKDG--WLAGDITATYDED 445
Qy	583	GFLRFHRTGDTFRWKGENVATTEVAEFEALDPELVNRYGVLPVPGEGRAGMAALVLR	642
Db	446	GFLRIVRKMKMILVSGFNVPENIEEVLMAHAKVLESVAIG--	VPEVSSGEAKVPEVVK 503
Qy	643	PPHALDMQLYTHVSENLPPARPRFLQESLATTFETFKQKVRMANE	691
Db	504	NDASLTPELTLTGRRYLTGYKVKIVEFRDELPSKNVNGILRRLEDE	552

RESULT 12

ID	LCRA_ECO57	STANDARD	FRT	561 AA.
AC	Q8XDR6			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	long-chain-fatty-acid-CoA ligase (EC 6.2.1.3) (long-chain acyl-CoA synthetase).			
GN	PAD OR OUND OR Z2848 OR ECS2514.			
OS	Escherichia coli O157:H7.			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;			
OC	Enterobacteriaceae; Escherichia.			
OX	NCBI_TaxID=83334;			
PN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=O157:H7 / EDL933 / ATCC 700927;			
RX	MEDLINE=21074935; PubMed=11206551;			
RA	Perna N.T., Plunkett G. Iii, Burland V., Mau B., Glasner J.D.,			
RA	Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,			
RA	Posfai G., Heckett J., Klink S., Boulton A., Shao Y., Miller L.,			
RA	Grotbeck E.J., Davis N.W., Lim A., Dimalanta B.T., Potoczniak K.,			
RA	Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,			
RA	Weich R.A., Blattner F.R.;			
RT	"Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";			
RL	Nature 409:529-533(2001).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	STRAIN=O157:H7 / RIMD 0509952;			
RX	MEDLINE=21156231; PubMed=11258796;			
RA	Hayashi T., Makino K., Okahishi M., Kurokawa K., Ishii K., Yokoyama K.,			
RA	Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tohe T.,			
RA	Kihara S., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,			
RA	Kihara S., Shiba T., Hattori M., Shinagawa H."			
RT <td>O157:H7 and genomic comparison with a laboratory strain K-12";</td> <td></td> <td></td> <td></td>	O157:H7 and genomic comparison with a laboratory strain K-12";			
RL	DNA Res. 8:11-22(2001).			
CC	-1- FUNCTION: Esterification, concomitant with transport, of exogenous			
CC	long-chain fatty acids into metabolically active CoA thioesters			
CC	for subsequent degradation or incorporation into phospholipids (By			
CC	similarity).			
CC	-1- CATALYTIC ACTIVITY: ATP + a long-chain carboxylic acid + CoA = AMP			
CC	+ diphosphate + an acyl-CoA.			
CC	-1- COFACTOR: Magnesium (By similarity).			
CC	-1- SUBUNIT: Homodimer (Probable).			
CC	-1- SUBCELLULAR LOCATION: Partially membrane-associated (Potential).			

CC -1- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
 CC family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC between the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE005403; AAG56794.1; -
 CC EMBL; AP002558; BAB35937.1; -
 CC PIR; B90943; B90943.
 CC PIR; F85791; F85791.
 CC InterPro; IPR000873; AMP-bind.
 CC Pfam; PF00455; AMP-BINDING; 1.
 CC PROSITE; PS00455; AMP-BINDING; 1.
 CC LIGase; Fatty acid metabolism; Magnesium; ATP-binding; Membrane;
 CC Complete proteome.
 CC NP BIND 213 224 ATP (PROBABLE).
 CC SEQUENCE 561 AA; 62364 MW; 4DE944AB7DF40CF2 CRC64;
 CC -----
 CC Query Match 7.1%; Score 273.5; DB 1; Length 561;
 CC Best Local Similarity 24.2%; Pred No 6.2e-11;
 CC Matches 125; Conservative 87; Mismatches 234; Indels 71; Gaps 18;
 CC
 CC 211 LSPGATVALLLPAGEEFLMFWGLAKAGRTAFVETALRQGLHCLSCGAPALVLAPE 270
 CC 71 LKKGGRVALMMENLQYPALEFGLRAGIVVNVNPLYPRELEHQLNDSASAIIVISN 130
 CC 271 FLESLEP-DLPA-----LRAMGLHMAAG-----PGTHPGISPLLAEV 308
 CC 131 FATHLEKVVDTAVQVHILTRMGDOLSTKGLTVNVVVKYIKRLVPEKTH---LPDAISFR 187
 CC 309 SAEVDGPVPGYLSPOSITDTCLYI-FTSGTGLFKPAARISHKTL-----OCQGFYQCG 363
 CC 188 SALHNGYRQGY-KPELVPEDLAFQYTGCTGVAKGMLTFRNLAMLEGVNATYGPIL 246
 CC 364 VHOEDVTYIALPLVHMSGSLGIVGCMGGA-TVVLKSKFSAGQWECQCHRYVFPQYI 422
 CC 247 HPGKEHVLVATPLPLVHIFALTINCLFIELGQGNLITFRDIPGLVKEKLAKYFPAITGV 306
 CC 423 GELCRVLVNPSPSKERGHKVLAVGSL--RPDTWERFVRFGGLQVLETYGLTE--GNV 479
 CC 307 NTLFNALLNKKERQQLDESSLHSAGCGPQVQVAERKVKLTGQ-YLLEGGLTECAPL 365
 CC 480 AITNTYGGAGAVGRASWIVKHLFPSSLIRYDVTTEPIRDPQGHCV-----ATSPGEFGL 534
 CC 366 VSVN-----PYDIDYHSGSIGLPVPSTPAKLVDDDDNEVSPQDGE 406
 CC 535 IYAPVQSQSPFLGYAGPELAGSKLLKDVFRGADVFPMTGDLVDDQGFRLFHDRTGDT 594
 CC 407 LC--VAGPQVLMGYWQRPD-ATDEITK-----NGMLHTGDIANDDEGFLRIYDRKDM 457
 CC 595 FRKKGENVATTEVAEFAEALDEPLQEVNVVYVTPVGHGAGAGAAVLVLRPHALDMQLYT 654
 CC 458 ILVSGNVVYVNEIEDVVMQHPGVQEVAAVGVP-SSSSGSAVKI FVYKDP-SLTRESLYT 515
 CC 655 HUSENTPYARPRFLQESLATEFFRQKRYRMANE 691
 CC 516 FCRRLQLGKVPKLVAFRDELPRXNVKILRELHDE 552
 CC
 CC RESULT 13
 CC LCFA_ECOLI STANDARD; PRT; 561 AA.
 CC ID LCFA_ECOLI
 CC AC P29212;
 CC DT 01-DEC-1992 (Rel. 24, Created)
 CC DT 01-FEB-1994 (Rel. 28, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE long-chain-fatty-acid-CoA ligase (EC 6.2.1.3) (Long-chain acyl-CoA
 CC synthetase).

GN FADD OR OLDL OR B1805 OR C2209.
 OS Escherichia coli, and
 OS Escherichia coli O6.
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; Escherichia.
 CC NCBI_TaxID=562, 217992;
 CC
 CC [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN=K12;
 CC MEDLINE=94150456; PubMed=8107670;
 CC RA Fulda M., Heinz E., Wolter F.P.;
 CC RT "The fadD gene of Escherichia coli K12 is located close to rnd at
 CC RT 39.6 min of the chromosomal map and is a new member of the
 CC RT AMP-binding protein family";
 CC RT Mol. Gen. Genet. 242:241-249 (1994).
 CC
 CC [2]
 CC SEQUENCE FROM N.A., AND SEQUENCE OF 1-10.
 CC STRAIN=K12;
 CC MEDLINE=93094273; PubMed=1460045;
 CC RA Black P.N., DiRusso C.C., Metzger A.K., Helmet T.L.;
 CC RT "Cloning, sequencing, and expression of the fadD gene of Escherichia
 CC RT coli encoding acyl coenzyme A synthetase";
 CC RT J. Biol. Chem. 267:25513-25520 (1992).
 CC
 CC [3]
 CC SEQUENCE FROM N.A.
 CC STRAIN=K12 / MG1655;
 CC MEDLINE=97426617; PubMed=9278503;
 CC RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 CC RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 CC RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 CC RA Mau B., Shao Y.;
 CC RT "The complete genome sequence of Escherichia coli K-12";
 CC RT Science 277:1234-1246 (1997).
 CC
 CC [4]
 CC SEQUENCE FROM N.A.
 CC STRAIN=K12;
 CC MEDLINE=97251358; PubMed=9097040;
 CC RA Itoh T., Alba B., Baba T., Fujita K., Hayashi K., Inada T.,
 CC RA Iono K., Kasai H., Kimura S., Kikawa M., Kikawa M.,
 CC RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
 CC RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T.,
 CC RA Saito N., Sempel G., Seki Y., Sivasubramanian S., Tagami H.,
 CC RA Takeda J., Takemoto K., Wada C., Yamamoto Y., Horichi T.;
 CC RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome
 CC RT corresponding to the 40.1-50.0 min region on the linkage map";
 CC RT DNA Res. 3:379-392 (1996).
 CC
 CC [5]
 CC SEQUENCE FROM N.A.
 CC STRAIN=O6:HI / CFT073 / ATCC 700928;
 CC MEDLINE=22388234; PubMed=12471157;
 CC RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
 CC RA Raeko D., Buckles E.L., Liu S.-R., Boutin A., Hackett J., Stroud D.,
 CC RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
 CC RA Mobley H.L.T., Donnenberg M.S., Blatner F.R.;
 CC RT "Extensive mosaic structure revealed by the complete genome sequence
 CC RT of uropathogenic Escherichia coli";
 CC RT Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
 CC
 CC [6]
 CC IDENTIFICATION OF ATP-BINDING MOTIF, AND MUTAGENESIS OF TYR-213;
 CC THR-214; GLY-216; THR-217; GLY-219; LYS-222 AND GLU-361.
 CC MEDLINE=22162442; PubMed=12034706;
 CC RA Weimar J.D., DiRusso C.C., Delio R., Black P.N.;
 CC RT "Functional role of fatty acyl-coenzyme A synthetase in the
 CC RT transmembrane movement and activation of exogenous long-chain fatty
 CC RT acids. Amino acid residues within the ATP/AMP signature motif of
 CC RT Escherichia coli FadD are required for enzyme activity and fatty acid
 CC RT transport";
 CC RT J. Biol. Chem. 277:29369-29376 (2002).
 CC
 CC -1- FUNCTION: Esterification, concomitant with transport, of exogenous
 CC long-chain fatty acids into metabolically active CoA thioesters
 CC for subsequent degradation or incorporation into phospholipids.
 CC -1- CATALYTIC ACTIVITY: ATP + a long-chain carboxylic acid + CoA = AMP
 CC + diphosphate + an acyl-CoA.

-1- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AEO08781; AAL20733.1; -
DR EMBL; AL627272; CAD05501.1; -
DR EMBL; AB016837; AAO68725.1; -
DR StyGene; SGR7777; faad.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 1.
DR PROSITE; PS00455; AMP BINDING; 1.
KW Ligase; Fatty acid metabolism; Magnesium; ATP-binding; Membrane;
KW Complete proteome.
FT NP_BIND 213 224 ATP (PROBABLE).
SQ SEQUENCE 561 AA; 62223 MW; 4000B2948FE3D15 CRC64;

Query Match 6.9%; Score 265.5; DB 1; Length 561;
Best Local Similarity 24.0%; Pred. No. 2.1e-10;
Matches 124; Conservative 80; Mismatches 243; Indels 69; Gaps 15;

QY 211 LSPGATVALLPAGPEFLMWFGKAKGRTAFVPTALRGGLHCLSCGAPLVLAPL 270
DB 71 LKKGRVALLMMPNLLQYVPAIFGILRAGIVVNVPLTYPRELEQLDMSGAALITSN 130
QY 271 PLESLEPDLP-----ALWAMGLHLMAG-----PCTHPAGISDLAEV 308
DB 131 PHTLEKVEKTSVGHVILTRMGDDLTAKGTAVNVFVKYIKRLVEXH--LPDAISFR 187
QY 309 SAEVDGPVGVGYSSQOSITDTCLTYFTSTTGLPKAAISHLKTL----CQOGFYQLGV 364
DB 188 SALQHGXYMQYKPEVVAEDLAFIQYTGTTGVAAGALTRHNMALNEQVATYGPILH 247
QY 365 HOEDVYIALPLPYHMSGSLIGVCGMGIGA-TVLSKFSAGQFEDCQCHRTVWFQYIG 423
DB 248 PKELELVTLPLPYHIFALTMNGLLPIELGGQNLTLTNRPDIPLGVLKELAKPFTAMGVN 307
QY 424 ELCRYVNVQPSKAERGHKVLAVSGL--RPDWERFVRRPGLQVLETGLTE-GNVA 480
DB 308 TLFNALNNKKEQQLDFSSILHSAGGMPQVNVVERWVKLTGQ-YLLEGGVLTLCAPLV 366
QY 481 TINYTGQRGAVGRASWLVYGHIFPSLIRYDVTGTGPIDPOCHM-----ATSPGEPGL 535
DB 367 SVN-----PHDIDHSGSIGLPVSTKXLYDDDDNEVAREGAGEL 407
QY 536 VAPVSGQSPDLGYAGPELAQGLKLDVFRPGDVFENTGDLVCDQGFLEFRHRTGDTF 595
DB 408 C--VKQSPVLMGYWGRPD-ATDEIRK-----DGMHTGDIAVMDEDFLRIVBRKKMI 458
QY 596 RKKGENVATTEVAEVEALDFLOEVNVGVYVPGHEGRAGMAALVLRPHALDLMQVLT 655
DB 459 LVSGFENVYNEIEDVMQSHGVQVEVAAGVP-SSGSGEYVKLFVVKDP-ALTDALITF 516
QY 656 VSENLPPYARPRFLRLQESLATEETFEKQKVMYANE 691
DB 517 CRRLHLYGVPRKOVERRELPRKSNVGLIRRELRLDE 552

RESULT 15
MENE_ENTFA STANDARD; FRT; 485 AA.
AC 0838K1;
DT 15-VAR-2004 (rel. 43, Created)
DT 15-VAR-2004 (rel. 43, Last sequence update)
DT 15-VAR-2004 (rel. 43, Last annotation update)
DE O-succinylbenzoate-CoA ligase (EC 6.2.1.26) (OSB-CoA synthetase) (O-succinylbenzoate-CoA synthetase).

GN MENE OR EF0446.
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX NCBI_TaxID=1351;

RP SEQUENCE FROM N.A.
RC STRAIN=V583 / ATCC 700802;
RX MEDLINE=22550857; PubMed=12663927;
RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,
RA Paulsen T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
RA Radd T.D., Dodson R.J., Umayam L., Brinkac L., Beaud M.,
RA Dagherty S., DeBoy R.T., Dalkin S., Kolonay J., Madupu R., Nelson W.,
RA Vamathevan J., Tran B., Upton J., Hansen T., Shetty J., Khouri H.,
RA Osterback T., Radune D., Ketchum K.A., Dougherty B.A., Fraser C.M.;
RT "Role of mobile DNA in the evolution of vancomycin-resistant
RT Enterococcus faecalis";

RL Science 299:2071-2074(2003).
CC -1- FUNCTION: Converts O-succinylbenzoate (OSB) to O-succinylbenzoate + CoA = AMP +
CC CoA (OSB-CoA).
CC -1- CATALYTIC ACTIVITY: ATP + O-succinylbenzoate + CoA = AMP +
CC diphosphate + O-succinylbenzoate + CoA.
CC -1- PATHWAY: Menadiquinone biosynthesis.
CC -1- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
CC family. Mene subfamily 1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AEO16948; AAO80301.1; -
DR TIGR; EF0446; -
DR HAMAP; MF 00731; -1 1.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 1.
DR PRINTS; PR00154; AMPBINDING.
DR TIGRFAMs; TIGR01923; mene; 1.
DR PROSITE; PS00455; AMP BINDING; 1.
KW Menadiquinone biosynthesis; Ligase; Complete proteome.
SQ SEQUENCE 485 AA; 54289 MW; 1BAD82BC8C08F82 CRC64;

Query Match 6.7%; Score 259; DB 1; Length 485;
Best Local Similarity 23.8%; Pred. No. 4.9e-10;
Matches 119; Conservative 71; Mismatches 213; Indels 96; Gaps 18;

QY 217 VALLPAGPEFLMWFGKAKGRTAFVPTALRGGLHCLSCGAPLVLAPLPELESLE 276
DB 51 VALFSKSKELVYSIALWELGKELLFATHLTLAELTQLDAQVKITIGAP----- 104
QY 277 PDLPALRMGLHLMAGPCTHPAGISDLAEVS-AEVDGPVGVYS-----SPOSTD 328
DB 105 -----TQALLLEISFVDVQPMIKKSHLSHQEFQSPSDLS 140
QY 329 TCIYFTSGTGLPKAARISHLK--LLQCGFYQLGVQGEQEVYIALPLPYHMSGSLIGI 386
DB 141 VASIMYTSGTGPQKAV-LQRFKNILASARGQENMGITAEBCMLCAVLPHIISG-LSLV 198
QY 387 VGCMGIGATVLLSKSPSAQFWDCCQHRVTVPQYIGELCRYLVNQPSPKAERGH--K 442
DB 199 VRQLTVGCSIRLYDKFDEQOVTDLQEGRGTVISVVAITMLQOLLISYVP--EAGYASAPK 255
QY 443 VRLAVSGRLRPDWEFRFRFGPLQVLEHYGLTE--GNVATINYQQRGAVGRASWLVYH 500
DB 256 GMLDGGPIAPDLACCEKGIIP-VISYGMETCSQVVALKFEEDALKITGSA----- 307
QY 501 IFFPSLIRYDVTGTGPIDPOCHCATSPGEGILVAPVSGQSP-FLG--YAGPELAQ 556
DB 308 -----GQLKXMQIK-----IYDELGGQGPKEQVGEIILKGNVVS 343
QY 557 GKLLKQVFRP---GVFVNTGDLVCDQGFLEFRHRTGDTFRKGENVATTEVAEVE 612

Db 344 GYL--NORQPEKMTADGWFKTGDMGYLDAQSYLYVSRSLSELISGGENITYPTEVEQYLQ 401
Qy 613 ALDPLQEVNVYGVTVPGHEGRAGMALVLRPPHALDLMQLYTHVSENI.PPYARPRFLRLQ 672
Db 402 AITGIKAAAVVG--EPDAQMGAVPVAVVI-SDQETTLAQIODCCSRKLAKYKRPKRIVFC 458
Qy 673 ESLATTEFFKQOKVMANE 691
Db 459 HSFQOTAGSKTAKRPMTE 477

Search completed: June 30, 2004, 17:33:53
Job time : 19 secs

This Page Blank (uspto)